THE UNIVERSITY OF ALBERTA

SYNTACTIC AND SEMANTIC VARIABLES IN THE LEXICON

by

© TERRY RIDPATH HULL

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF Doctor of Philosophy

IN

LINGUISTICS

DEPARTMENT OF LINGUISTICS

EDMONTON, ALBERTA

FALL, 1978

For Kerry and Christopher

ABSTRACT

The effects of selected syntactic and semantic variables on the learning of paired associates were investigated in an attempt to address the larger problem of the nature of chunking in language processing. The three syntactic categories studied were nouns, adjectives, and verbs, and the four semantic properties were frequency, rated meaningfulness, scaled imagery, and scaled abstractness.

Subjects were required to learn lists of real word-nonsense syllable paired-associates. The correct responses were tallied and converted to the learnability scores used in the subsequent analyses. The results showed that syntactic facilitation was operative in the learning of the adjectives. It was impossible, however, to confirm whether or not it was operative for the nouns and verbs. In contrast to earlier studies, the present study found that meaningfulness exerted greater influence than any of the other semantic variables and not imagery as has been suggested by numerous researchers. The author concluded that this discrepancy could be attributed to the failure of earlier studies to control for the two variables simultaneously.

The data suggest that the relationship between the

syntactic and semantic variables is both intricate and varied. While it is impossible to claim the clear superiority of either syntax or semantics in the chunking process, it can be established that bonding entails the use of information from both areas. It seems to be the case that even a strong syntactic factor can be overridden, in certain circumstances, by a powerful semantic factor, and, conversely, that a strong semantic factor can be overridden by a sufficiently well established syntactic factor.

ACKNOWLEDGEMENTS

I would like to express my gratitude to the members of my committee: to Dr. John Hogan for his help and encouragement with the statistical analysis; to Dr. Lois Marckworth for her constant support and excellent editorial advice; to Dr. W. Runquist for his suggestions and thought-provoking comments at candidacy and defense; to Dr. Gary Prideaux for his assistance during the final stages of thesis preparation, and, in particular, to Professor W.J. Baker for his guidance and assistance at every stage of my Ph.D. program.

I would also like to thank Dr. Anton Rozypal for his invaluable assistance with computer programming for the experiments, and Richard Stephens for his help with the execution of the experiment. The graduate students of the Department of Linguistics deserve special thanks for their participation in the pilot experiments and for their advice and criticism throughout. To Dale Stevenson and Grant Crawford, I owe a large debt for their patience and their aid in production of the final manuscript.

The research was made possible, in part, by a departmental research grant, and by graduate teaching assistantships. I wish to thank the Department of

Linguistics for this support.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

This study is intended to provide partial answers or, at least, clues to some fundamental questions in psycholinguistics. Foremost among these is the question of semantic organization, or the properties of the lexicon. In recent years, research on that problem by linguists, though claimed to be psycholinguistic, has been mostly formal. Research into semantic networks, for example, has proceeded from the desire to incorporate semantics into a generative-transformational grammar. There has been very little interest, among formal linguists anyway, in investigating the nature of the speaker/hearer's internal, psychological semantic network. The reasons for this attitude are clear when it is noted that in recent years the term "psycholinguistics" has been synonomous, for most linguists and even some psychologists, with transformational-generative grammar (TGG) (cf. Fodor, Bever, and Garrett, 1974; Cairns and Cairns, 1976). Since historically the development of TGG has been based on syntax, the central question with regard to semantics has usually been its relationship to syntax in a formal grammar. Still the claim has been that TGG is psycholinguistics, that it is psychologically viable. That

these claims could be made without empirical support of any kind astounds many psychologists; yet that is exactly what has been done.

While linguists have clearly overemphasized the role played by syntax at all levels of language use, psychologists have often failed to give it enough consideration. Simply at the level of the lexicon, when a person "knows a word," he knows far more than its reference. He has attitudes toward its referent; he has, idiosyncratically, experiences, images, and associations with respect to it. Much of this psychology has, at various times, attempted to quantify. In addition, though, the language learner also knows a great deal about the role or roles a term can play in his linguistic system. It has been assumed by many experimenters, but particularly by linguists, that this knowledge is syntactic in the sense that the speaker/hearer "knows" about syntactic categories and syntactic roles, that he knows, for example, that bartender, John, doctor and woman are all nouns and that pour, walk, practice and work are all verbs. It is likely that the speaker/hearer does have some basis on which to classify these two groups of words, that, indeed, he is aware of the roles they play in larger communication units. The experimental evidence available tends to support this view. But it must be emphasized that to assume that this knowledge is syntactic is unwarranted. While it is possible

to assign syntactic labels (and rules for that matter) to the language product, there is no reason to assume that these correspond to people's knowledge. It is entirely possible that what they are aware of are semantic attributes, e.g., agent and action that can be described in non-syntactic terms, e.g., as case relations. In other words, syntax may be an externally imposed notion on relationships and attributes that are primarily semantic. And all of these variables can, in as yet unclearly specified ways, act to influence how a given word will function, either as a stimulus or as a response, in empirical studies of any kind.

Some sort of "semantic network" is presumed by most cognitive psychologists, to be a necessary organizational component of memory. However, the problem of defining it is easily seen by considering what is involved in retrieving something stored in memory. Most of the time it is a quick, efficient process. But when there is difficulty, that difficulty reveals much of the complexity of the memory structure. A person trying to recall the word barrister, for example, might come up with various "clues," i.e., he might search various levels and come up with British, lawyer, court, client, bear, or just b. He would probably not, however, ever come up with noun, subject, or object although he would not likely go outside the category that is commonly called noun, suggesting that there are category constraints

of some sort operating. But, of course, the constraints are more rigid than just category constraints as he would not produce nurse, colonel, or cleaning woman even though they fit both the syntactic category, noun, and the case category, agent. The point is that he is searching different levels of his memory, or, alternatively, trying different routes of access.

The present study was undertaken in an attempt to explore certain of these routes. Specifically, it was intended to discover the extent to which subjects take advantage of syntactic information and which semantic parameters provide the best access routes -- in other words, which are psychologically viable.

Background

Since psychologists and linguists began talking to each other again, in the 1950's, there have been only a few areas on which they have clearly understood each other and agreed. One of these is the existence of language units larger than the word. They have not agreed, however, on the nature of these units. Linguists have talked of constituent structure to refer to the hierarchical structure of a sentence, which they demonstrated in page after page of impressive-looking trees. The psychologists have had something rather different in mind, referring to the human

processor's organizing or grouping of bits of information into larger chunks to reduce the load placed on immediate memory. Thus, the digital sequence 384961475 might be recalled as three chunks, 384, 961, and 475, rather than as nine separate integers.

The chunk was introduced to psychology by George Miller in 1956, and since then this concept of recoding has been vitally important to research on information processing. A powerful recoding system permits the human processor to store large amounts of information in a clearly limited memory system. Miller believed that language could be considered as a predominate recoding system. Citing Hayes (1952) finding that the memory span for English monosyllabic words is five, Miller noted that while it might be a logical claim that the subjects actually recalled approximately 15 phonemes, it is intuitively obvious that what they recalled was five words. This suggests that the subjects may have recoded the information from a lower level into higher level chunks that were more readily recalled. It is possible, of course, that the five words were <u>perceived</u> directly as five words even though they could be analyzed into phonemes, phones, features, etc. Recoding of words into higher order "chunks," however, clearly takes place in processing complex utterances. While it is possible to do recoding in many ways, "probably the simplest is to group the input events, apply a new name to the group, and then remember the new

name rather than the original input events" (Miller, 1956, p. 93).

As a consequence partly of Miller's work and partly of Chomsky's (1957, 1963, 1965), psychologists and linguists have spent much of the past two decades investigating the nature of the constituent or chunk. More accurately, the linguists have tried to establish psychological reality for the hierarchically arranged descriptions of the language product, i.e., phrase-markers. As Martin (1970) pointed out, the particularly salient thing about phrase-markers is their invariance over authors. Most would agree, for example, that the structure of Mary slapped John was ((Mary) (slapped (John))). This traditional descriptivism underlay the click studies of Fodor and Bever (1965) and Garrett, Bever, and Fodor (1966), the transition-error studies of Johnson (1965), and the Yngve depth studies of Martin and Roberts (1966). All of these investigators began with TGG deep structure phrase-markers and looked for their representation in signal detection, in sequential learning, and in sentential recall. Levelt (1969) quotes Rommetveit's observation that the results of these experiments are "supportive of any theory that assigns some perceptual reality to large syntactic constituents" (Levelt, 1969a, p.4).

Because results from such studies have been inconclusive, Martin rightly asserts that it is "imperative

to discover empirically whether or not our time-honored, widely accepted views on phrase-structure are in fact representative of how language users themselves organize sentences" (1970, p. 153). Levelt recognized the problem as well:

> We do know that 'chunking' is taking place during perception of sentences, but we know insufficiently about the linguistic nature of such chunks. The idea that they correspond to the surface constituent structure of the sentence as specified by transformational grammar can neither be accepted nor rejected at our present stage of data analysis. (Levelt, 1969a, p.4)

Levelt's concern with what Martin (1970) called subjective phrase-structure, and its relationship to traditional descriptive phrase-structure, led him to conduct a series of experiments which will be reviewed briefly here. The first experiment, reported in 1969 (and in Levelt, 1970), was concerned primarily with validating Johnson's (1967) hierarchical clustering scheme as a technique for demonstrating how people intuit syntactic relations and how they chunk information in sentence processing. The experiment employed 20 sentences of various syntactic structures, played through white noise to 120 subjects. The subjects had been instructed to write down all words they were able to identify. For each of the eleven sentences found to have an error rate in the range of 30 to 70%, an n x n matrix of conditional probabilities was computed. For any two words, x and y, from a sentence, Levelt calculated

the probability that y had been correctly identified, given correct identification of x, and conversely that x had been identified given the correct identification of y. Conditional probabilities were computed for all pairs of words from each sentence.

Johnson's hierachical clustering analysis was applied to the eleven sentences to discover whether a left-to-right hierarchical chunking model would fit the data.

> Each clustering specifies a set of [psychological] distances between words. If the model fits the data the rank order of these distances should be inversely related to the rank order of conditional probabilities in the data matrix. (Levelt, 1969a, p. 5)

Levelt found the model, on the whole, adequate to describe the data. In other words, using this technique he was able to confirm the subjects' hierachical chunking of sentential information. Furthermore, he found that subjective phrase-structure corresponded to linguistic, i.e., descriptive, phrase-structure, in the case of main constituents but not in the case of minor constituents. For the sentence, the tall boy saved a dying woman, Levelt was able to confirm only the two main constituents, (the tall boy) and (saved a dying woman).

It would be faulty to conclude, however, that the adjective-noun construction, dying woman, did not have "unit" or "constituent" status. What is more likely is that

Levelt's technique was insensitive to the minor constituent distinct from the major verb phrase constituent in which it occurred.

For another sentence, the house of the baker is on fire, Levelt found an "ungrammatical" clustering of of the. Even though this clustering is unlikely from a purely descriptive point of view, Levelt believed it to be perfectly natural subjectively since many languages contract such word pairs.

A second series of experiments (1969b) applied Johnson's hierarchical clustering scheme to subjects' judgments of the degree of relatedness between the words of a sentence. In one experiment, subjects were presented with the sentence the boy has lost a dollar, and instructed to judge the degrees of relatedness between the words of that sentence. The comparisons were triadic, i.e., subjects were given a set of drawn triangles with words from the sentences at the vertices, and told to mark the pair which was most closely related with a "+" and the pair showing the least relationship with a "-".

Levelt computed for each word pair the relative frequency (over triads and subjects) with which it was judged as more related than other pairs. He then applied Johnson's method to the data to demonstrate the subjects' clustering of the words. He found that the resulting

hierarchical structure corresponded closely with the descriptive phrase-structure of the sentence.

A second experiment required Dutch subjects to make relatedness judgments for the sentence, Jan eet¹ appels en Piet eet² peren (John eats apples and Peter eats pears). This task, however, was a rating task. Each subject was given the sentence and a list of all the pairs of words from the sentence accompanied by a 7-point rating scale for each pair. The clustering analysis applied to these data once more revealed a high correspondence between subjective and descriptive phrase-structure. Levelt was careful to point out that phrase-structure is not the sole determinant of relatedness judgments: "Words similar in meaning and adjacent words tend to have slightly inflated relatedness values" (p. 352).

Levelt's conclusions, qualified as they are, must be viewed with some skepticism. There are three problems. In the first place, the rating technique used by Levelt might yield uninterpretable results with regard to discontinuous constituents.

> To the extent that rating judgments are influenced by ancillary associative and semantic considerations, non-adjacent words may be overrated in relatedness, thus increasing the likelihood of an unintended discontinuous constituent. (Martin, 1970, p. 165)

Martin's point is important. What he implies is that

Levelt may have misinterpreted his results by assuming that the subjects' chunking techniques were syntactically based when, in fact, they were semantically based. In other words, the strong bond between eet and peren in Jan eet appels en Piet peren may have nothing to do with syntactic structure at all. While it cannot be argued that there is no syntactic relationship, for certainly there is, it may be the case that the subjects are unaware of it, that their linking of the two words is done on the basis of the meaning of the sentence: they simply know that what happens to pears is that they are eaten. It is more likely, of course, that syntactic and semantic information together account for the strength of the bond.

The second problem is one not uncommon to psycholinguistic experimentation. The problem is with comparing sentence types which are unequal in their probability of occurrence. In English, for example, the "unreduced" form (e.g., Carla takes her books and Carla goes to school) is far less common than the "reduced" form (Carla takes her books and goes to school). In the experimental situation, the former sentence presents a task which is far less natural than a task involving the latter sentence. Furthermore, there is a presupposition operative here which is characteristic of experimentation which has set out to establish the psychological salience of TGG. The presupposition is that the "reduced" form is derived from

the "unreduced" form. From a descriptive point of view, there is no objection. It is a mistake, though, to assume that this is necessarily the relationship between the two sentences in any psychological sense. It is absurdly counter-intuitive, in fact, to assume that a speaker "generates" two identical subjects and then "reduces" them to one. Indeed, there would be as much reason to assume that, from a processing point of view, the "base" form is the reduced form since it occurs with far greater frequency in the language, or since children learn it earlier. The point is that any assumption of a psychological "base" form is spurious without supporting empirical evidence.

The third problem with the Levelt studies is in generalizing his results. He dealt, after all, with only a few sentences, one for each experiment, and all of these had fundamentally the same Subject-Verb-Object (SVO) structure. There was no attempt made to investigate the correspondence between subjective and descriptive phrase-structure with other types of sentences, e.g., Subject-Verb (SV) or Subject-Verb-Adjective S-V-ADJ). Moreover, all the sentences used were short and simple. Possibly, only more complex types could demonstrate differences that could be interpreted as corresponding to minor constituents. Certainly there is no reason to assume that the processing of short simple sentences and the processing of longer complex sentences can be equated even though they may be

related in a syntactic description. The main point is that it would be a mistake to generalize Levelt's conclusions to the perceptual organization of all sentences, and Martin's (1970) findings make this point clear.

Martin applied a slightly modified version of Johnson's hierarchical clustering analysis to the data obtained from 60 subjects sorting the words from 45 different sentences into clusters. There were 15 sentences each of the lengths 7, 10, and 15 words. Within each length were three different sentence patterns with five lexically distinct sentences per pattern. The subjects were instructed to arrange the words of each sentence into groups with no restrictions. In other words, the group could be one word, the entire sentence, or anything in between, and the subjects had no time limit.

Martin found that for sentences of the pattern Subject-Auxiliary-Main Verb-the-Adjective^1-Adjective^2-Object (e.g., Parents were assisting the advanced teenage pupils), the major subjective constituents were not (S(VO)) as TGG descriptive phrase-structure would predict, but ((SV)O). The same hierarchical method applied to sentences of the pattern S-Pronoun-V-Adverb^1-MV-O-Adverb^2 (e.g., Children who attend regularly appreciate lessons greatly) yielded a different clustering: (S(VO)).

Martin's and Levelt's studies report results which are somewhat contradictory, but they are highly significant in

one respect: they represent an important first step in the right direction. That is, they make it clear that an investigation of the language process, whether it be perception or production, must begin with the language user and not with the language product. There is, after all, no a priori reason for assuming that any particular description of the structure of the language product should correspond to the structure assigned the sentence by the user. Martin's and Levelt's studies recognize that the link between subjective and descriptive structure is not as strong as earlier (transformationally based) research would have us believe.

Martin and Levelt clarified the problem and established the direction that research must take. They provided valuable hints, but they did not discover the basis on which hierarchical clusters are formed and then re-formed into larger chunks. The specific problem is to identify the properties of words that cause them to bond together to form these chunks. The next section of this paper reviews a series of studies which provide some further clues. Most of these studies were concerned not with verifying the psychological salience of descriptive phrase-structure, but with discovering the effect of syntactic structure on verbal learning.

The Syntactic Parameters

Epstein (1961) investigated the effects of syntactic clues on the learning of nonsense syllables. He chose nonsense syllables in order to control factors such as meaningfulness, familiarity, and sequential probability. The nonsense syllables were arranged in "sentences" in six categories.

The first consisted of two sentences composed of nonsense syllables and two function words without referents i.e., articles, conjunctions, or prepositions. The nonsense syllable-stems were suffixed with grammatical markings such as -s on "plural nouns" and -ed on "verbs" (e.g., A vapy koobs desaked the citar molently um glox nerfs). The second category was the same as the first except that the grammatical tags were omitted (e.g., a vap koob desak the citar molent um glox nerf).

Category III had the same material as Category I, but the items were presented in random order (e.g., yigs rixing wur miv hum vumly the jegest). The fourth category retained both the syllable-stems and the order of the first category. The positions of the grammatical tags were shifted in an attempt to induce the formation of a competing syntactical pattern, one incongruent with customary English usage (e.g., The yigly wur vums rixest hum in jeging miv.).

The fifth and sixth categories consisted of meaningful words arranged in an order that was semantically anomalous and within which existed a low level of transitional probability between neighboring words. The fifth category contained two series of words so ordered as to be syntacticlly acceptable but semantically anomalous (e.g., Cruel tables sang falling circles to empty bitter pencils.) while Category VI was both syntactically and semantically anomalous (e.g., sang tables bitter empty cruel to circles pencils falling).

Subjects were instructed to learn each string of syllables in the order presented. Each string was presented individually for seven seconds after which the subject was given 30 seconds to write his response. If the response given was incorrect in any way, the experimenter presented the string again, and continued to present it until the correct response was given. Once criterion learning was reached for one string, the experimenter moved on to the next.

Epstein found that Category I required significantly fewer trials than Categories II, III, or V, and that V required fewer trials than VI. The difference between Category I and Category IV was statistically non-significant. Attempting to explain his results, Epstein hypothesized that "material which is not syntactically

structured may be harder to learn than structured material because the latter is already organized whereas the former must be organized into more efficient chunks only through the intentional efforts of the learner" (p. 84). In other words, the learner takes advantage of syntactic structure if it is available to him. If it is not, he will, presumably, attempt to impose his own structure, i.e., he will organize the information into chunks in order to facilitate recall.

Wind and Davidson (1969) attempted to determine whether paired-associate (PA) learning of word-nonsense syllable pairs was enhanced by providing semantic and syntactic context for the nonsense syllables. The investigators employed two learning conditions. One was the traditional PA frame in which subjects saw only the word-nonsense syllable pair, e.g., BAP-DINNER. The other was the context learning condition in which subjects saw the nonsense syllable embedded in English sentences with the real words following in parentheses, e.g., "What time is BAP served? (DINNER)" (p. 185).

Following the learning trials, subjects were given two tests under two different testing conditions. The first was the PA test in which subjects were required to give a nonsense syllable response to an English word stimulus. The second was the context test in which subjects were required to fill in the blank spaces in a paragraph with the appropriate nonsense syllables. "The words omitted were

those used in the learning phase, and the context was redundant, in the sense that the English words could be easily determined from the context" (Wind and Davidson, 1969, p. 185). In both tests all of the nonsense syllables used in the task were listed at the bottom of the page, thus reducing the problem of response learning and making the task more one of recognition than of recall. Both the real words and the nonsense syllables were controlled for meaningfulness, all having high meaningfulness or association values.

The results of Wind and Davidson's experiments showed that of the two learning conditions, subjects learned better when the pairs to be learned were embedded in English sentences. Of the two recall tasks, however, subjects did better when only the stimulus word was given (no real-word context). The investigators suggested that this second finding may have occurred "simply because the English words were not seen easily enough in the blanks of the paragraph presented" (p. 185). What is equally likely is that the context merely got in the way and interfered with correct responses once they had been learned. It is the first finding, however, that is significant. It provides evidence that the learning of associations between real English words and nonsense syllables is facilitated by syntactic context. In other words, there is further support for Epstein's hypothesis that syntactically structured material is easier

to learn because it is already organized into efficient chunks.

Glanzer (1962) reported findings consistent with those of Epstein and Wind and Davidson. In a series of three experiments he attempted to demonstrate that content words, i.e., nouns, verbs, adjectives, and adverbs, were more easily learned than function words, i.e., prepositions, conjunctions, and pronouns, in the paired-associate paradigm. Indeed, his results have been reported as conclusive evidence that it is easier to learn associations between nonsense syllables and content words than between nonsense syllables and function words (Miller, 1962; Bacharach, Kellas, and McFarland, 1972; Fodor, Bever, and Garrett, 1974).

Glanzer's technique in the first two experiments was to pair words from each of the seven grammatical categories listed above with Consonant-Vowel-Consonant (CVC) nonsense syllables, positioning half the words before the nonsense syllables and half following. He interprets his findings to mean that "in general . . . the content words are learned more readily . . . " (p. 33).

Baker and McCarthy (1975), however, pointed out that Glanzer might have misinterpreted his results. In the first place, he went "directly to a test between his two a priori, non-empirical, classes without first establishing whether or

not it was valid to assume a generic class effect" (Baker and McCarthy, 1975, p. 6). Glanzer's graphs (p. 33) show that there are within class differences, e.g., nouns and adjectives are more easily learned than any other category while verbs and adverbs differ little from the function words.

Furthermore, the analysis showed that all word types were learned more easily when they followed the nonsense items than when they preceded them. Had any kind of syntactic facilitation occurred in the learning of the word lists, some degree of differentiation would have been expected. That is, because the different word types all have different syntactic functions, it would be reasonable to expect some position-by-type interaction. Glanzer offered no explanation for his "position effect," but the usual explanation for an effect of this type is that syntactic information facilitates learning. In other words, knowledge of the syntactic role of the various word types induces a surface constituent structure onto the word-nonsense syllable pair. Consequently, the unrelated items become related, or bonded, and thus more learnable. Indeed, this is exactly the explanation Glanzer invoked in his discussion of his third experiment.

Embedding real words between nonsense syllables, Glanzer found that subjects learned the triplets with function words more quickly than triplets with content

words. His explanation was that placing the function words between the syllables permitted them to operate as they normally would - as conjunctions (mev and bep or bread and butter) or prepositions (mig of gav or cream of wheat). Glanzer's explanation for these results is reasonable and likely correct. But any such explanation of the position effect in the first two experiments is doubtful.

In fact, Glanzer's position effect is an artifact of his experimental technique. For half the data, Glanzer's subjects had to recall nonsense forms as responses; for the other half they had to recall meaningful words as responses. "Any explanation for his results suggesting entailment of syntactic facilitation is clearly pointless given his methodology" (Baker and McCarthy, 1975, p. 8). Had syntactic facilitation been operative in the experiment, it could not have been revealed using Glanzer's design since the tasks - recalling a nonsense item and recalling a real word - are not comparable. Real words are probably easier to remember, but because of their inherent richness in meaning relative to nonsense items, there is no assurance whatever that their syntactic category contributes to this ease of recall. Still, syntactic facilitation is the implied explanation both in the Glanzer study and in citations of it. Epstein (1961) demonstrated that knowledge of the syntactic role of a particular word would induce a surface constituent structure making previously unrelated items

easier to learn. Because of the position effect, it is doubtful that this kind of syntactic facilitation occurred in the Glanzer study. That this is true becomes obvious when we consider how each word type is likely to function. For example, a noun preceding a nonsense syllable might be linked or bonded to the latter by considering the nonsense syllable as an unfamiliar intransitive verb, or as part of a compound noun. The reverse position, CVC-noun, might be perceived as a noun phrase consisting of an adjective and a noun. Baker and McCarthy (1975) point out that it would be difficult to predict which type of constituent would be easier to learn, so some degree of syntactic facilitation would be a reasonable expectation in either position.

For adjectives, on the other hand, there should be a clear position preference for English speakers. Placing an adjective before a noun is likely to result in the perception of a noun-phrase constituent. Glanzer's data, however, do not support this hypothesis. Rather, they show that learning occurs faster when the adjective follows the nonsense syllable.

For verbs in the Glanzer study, it would be expected that the position following the nonsense syllable might be preferred. In this position the pair might be construed as an imperative construction, e.g., Think-Mef. Again, the opposite results were obtained. The preference for the CVC-verb pattern can hardly be explained by syntactic

facilitation since the absence of either third person singular morphology for the verb forms or plural morphology for the noun forms makes subject-verb bonding unlikely unless we assume that subjects construed the nonsense forms as irregular noun plurals.

Obviously, Glanzer's position effect cannot be explained by syntactic facilitation. More likely, the effect is a result of his technique.

> For subjects learning the word-nonsense syllable sequence, the response was . . . a meaningless item. For the other subjects, it was a meaningful word. This would appear to be the obvious reason why the latter task was so much and so consistently easier. (Baker and McCarthy, 1975, p. 13)

Kanungo (1969) extended Glanzer's study, replicating both the results and, unfortunately, the errors in interpretation. His results showed greater learnability for CVC-word order than for word-CVC order. Furthermore, he found that function words embedded between CVC syllables were easier to learn than content words in the same position. He did, however, make the same mistake as Glanzer in not checking for within-class differences. He merely assumed a generic class effect.

Baker and McCarthy corrected the major weaknesses in the Glanzer study (and in the Kanungo study) in their experiment investigating syntactic facilitation. A major change from the Glanzer study was to use only real words as

stimuli in order to impose as few restrictions on the stimuli as possible. These words were paired with nonsense syllables in six lists, one each for adjectives, non -ly adverbs, -ly adverbs, nouns, verbs, and prepositions. In each list half the real words preceded their associated nonsense items and half followed them, but real words were always used as prompts and nonsense items as responses.

The experimenters presented each item on a memory drum in a three-stage display as follows:

(1)	*	*
(2)	stimulus	*
(3)	stimulus	response

or

(1)	*	*
(2)	*	stimulus
(3)	response	stimulus

(The motivation for this type of presentation will be discussed in a later section of this paper.) Half the subjects were given the three modifier lists: adjectives and the two types of adverbs, while the other half were given the prepositions, nouns, and verbs. The lists were presented until the subject could give all correct responses for two

consecutive blocks of items.

For the adjectives and adverbs, the analysis of the data showed a significant interaction between word types (adjective or adverb) and position (preceding or following the nonsense form). This finding is in marked contrast to Glanzer's findings discussed earlier. For the real words located after the nonsense items, there was little difference among the word types. When the real word preceded the CVC, however, "the adjectives become much more effective as stimuli, suggesting the kind of syntactic facilitation that ought to occur with English in the common formation of a noun phrase" (Baker and McCarthy, 1975, p. 20).

Baker and McCarthy also found that adjectives were, in general, easier to learn than either type of adverb, a finding consistent with Glanzer's. But, of course, Glanzer and citations of his study ignored this fact in reaching the conclusion that content words are, in general, more easily learned than function words.

The data for Baker and McCarthy's second group of subjects included nouns, verbs, and prepositions. The analysis indicated that position before or after the nonsense form affected the learnability of the verbs but not the prepositions or nouns. Apparently it was easier for subjects to form constituents resembling verb phrases (Gave-

Jat, Wrote-Kex) than it was for them to form a subject-verb constituent. The investigators offer the explanation that the former is a relatively complete constituent while the latter, for transitive verbs, is not.

Bacharach, Kellas, and McFarland (1972) found no significant difference between pre- and post-position of the CVC in their study investigating the differential learnability of transitive and intransitive verbs. Theirs, however, was a free recall task (no prompt was given), so no direct comparisons with the Glanzer, Kanungo, or Baker and McCarthy studies are possible. (The investigators did find evidence to support their hypothesis that subjects make covert noun responses to complete incomplete transitive verbs - intransitive verbs were easier to learn than transitive in the CVC-verb position.) It would be reasonable to expect that the reverse position would yield reverse results, that transitive verbs would be easier to learn in the verb-CVC position since an "overt" response, or a complete syntactic unit would be present. This was not the case: there was no significant difference between transitive and intransitive verbs in the post-position. These findings might be taken as evidence that S-V has more psychological import as a constituent than does V-O, but more likely they merely support Rohrman's (1968,1970) finding that intransitive verbs are, in general, easier to learn than transitive verbs.

Prepositions were learned equally well, in the Baker and McCarthy study, when placed before or after the nonsense items, suggesting a prepositional phrase type constituent in the first case and suggesting a verb-particle constituent in the second. Similarly, nouns were seen equally well as noun modifiers, when they preceded nonsense items, and as head nouns in a noun phrase, when they followed nonsense items. This fact may contribute to the finding that noun lists were learned more rapidly than either preposition or verb lists. But since prepositions showed the same lack of position effect, there must be other contributing factors. Baker and McCarthy suggest that one of these is the higher degree of meaningfulness (as Noble uses the term) and association value of nouns.

Summarizing Baker and McCarthy's findings, it is clear that (a) the existence of a "content" class is doubtful given the differences between the members of the class, so any claim for differences between classes is vacuous, and (b) syntactic facilitation operates for adjectives and verbs as it would be expected to under the appropriate experimental conditions. While there are certainly differences between word types that must be taken into account in learning tasks involving these types, it would be a mistake to pool data across types. As Baker and McCarthy put it:

> It is evident that some form of syntactic facilitation occurs producing differential effects for verbs and adjectives (or, at least, for the limited set of exemplars examined thus far). This may be inoperative or, more likely, simply have an equally facilitative effect for other word types in either position. Overall type differences are probably related to the more traditional parameters in this area (Cofer, 1971): m, m', association value, etc. , but these must be explored systematically for the various word types. (1975, p. 23)

That some semantic factor or factors might influence the PA learning task more than any syntactic considerations was suggested by Paivio's (1963) study using adjective/noun pairs. In a preliminary study, he had found that subjects learned the noun-adjective order more easily than the adjective-noun order, a finding contrary to expectation considering the structure of the English noun phrase. He hypothesized that "nouns act as conceptual 'pegs' from which their modifiers can be hung." Thus, nouns served more effectively than adjectives as stimuli. He further hypothesized that concrete nouns should be more "solid" pegs than abstract nouns and tested this hypothesis in his first experiment. The results of that experiment indicated that subjects once again learned the noun-adjective order better (i.e., with fewer errors) than the adjective-noun order.

Furthermore, scores for concrete nouns were higher than for abstract with either order, but scores for abstract nouns in the N-A order were higher than for concrete nouns

in the A-N order. Paivio attributed this latter result and the fact that the differential effect was slight to the fact that the selection of the nouns for the test had been too highly restricted because his subjects were children. In other words, only small differences in abstractness were possible using the high frequency nouns appropriate to children. "Although the 'abstract' nouns are more general than the 'concrete' nouns, many also have concrete referents" (Paivio, 1963, p. 374). Therefore, the study was repeated using university students as subjects and words of a frequency level permitting the selection of more highly differentiated abstract and concrete nouns.

The results of the second experiment showed once more that the effects of word-order and noun abstractness were highly significant. Recall was better for N-A order than for A-N order, and better for pairs with concrete nouns than for pairs with abstract nouns. Paivio offered as an explanation of the word order effect the possibility "that adjectives are more frequent associates of nouns than <u>vice versa</u>, despite English grammatical habits" 1963, p. 377). In other words, in the PA task it was easier for subjects to recall the adjectives which were paired with the nouns than it was to recall the nouns which were paired with the adjectives. Paivio interprets his results as consistent with the conceptual peg hypothesis discussed above. There are, however, several problems with Paivio's study which

cast doubt on his interpretation.

One problem is that the experimental procedure was "loaded" in favor of the conceptual peg hypothesis. There was no reason to expect, given his methodology, that syntactic facilitation would be operative. The subjects in both experiments "learned" the paired associates via a single oral presentation and were subsequently required to write responses to oral stimuli, half of which were nouns and half of which were adjectives. It is doubtful whether this procedure would permit subjects to take advantage of syntactic information. The results can be explained by the fact that a noun simply has a more limited number of possible adjective associates than an adjective has noun associates. Using exemplars from Paivio's first (1963) study, it can be seen that there are fewer possible adjective associates to engine and glass than there are noun associates to white and flat. The conceptual peg hypothesis is weakened, then, if it is the case that nouns are superior stimuli only because there are fewer possible adjective responses to be confused with the correct one (as may have been the case in Paivio's study).

Paivio's hypothesis is weakened even more when the items used in the two experiments are examined. There was no attempt made to control for pre-existing associations. Consequently, the items used in the first experiment included proud-Indian, open-window, hungry-bear, and soft-

chair in the concrete lists and the same adjectives paired, respectively, with people, space, animal, and spot in the abstract lists. The second experiment included technical-advertisement, clumsy-burglar, colourful-maple, and heroic detective in the concrete lists and the same adjectives paired with discourse, imitation, scenery, and destiny in the abstract lists. Many of these pairs are so frequently and uniquely paired in the language that their status approaches that of a compound noun. This fact underlines the seriousness of the objection made above to Paivio's conclusions. Especially for word pairs that are closely associated in the language, the noun will serve as a better elicitor for the pair simply because there are fewer possible meaningful adjectives available for association than there are nouns for adjective stimuli. As Keenan pointed out, it is the noun which "selects" the adjective, the possibilities from which to select being fewer than in the reverse case. Furthermore, adjectives are more variable in their meanings than are nouns, and in fact, the meaning of an adjective is somewhat "conditioned" by the noun with which it is paired. The meaning of strong, for example, differs when it is paired with table from when it is paired with tea, man, or argument.

Two additional objections to Paivio's study should be raised. Both deal with his claims about abstractness. The choices of items for his studies were not based on rated

abstractness values as were his later studies, but merely on a definition of abstract words as those having no objective referents. Studies since (Paivio, 1968; Baker and Hull, 1976) have shown that there are degrees of abstractness, i.e., that the relationship between the two attributes is not clearly dichotomous, a point which will be discussed more fully in the next section of this chapter.

The second objection regarding abstractness concerns Paivio's failure to distinguish the general-specific and abstract-concrete dimensions. Noise, which he treats as abstract is more general, certainly, but not necessarily more abstract than shot. Similarly, scenery is more general than maple, but not necessarily more abstract. Since it cannot be said with certainity that Paivio even differentiated the abstract from the concrete nouns, it is difficult to trust any conclusions about the superiority of one over the other in his learning task.

The final criticism of this study is perhaps the most serious. It concerns the failure to control for imagery (the capacity a word has for eliciting mental images). A subsequent study by Paivio (1969) and others reported later in this chapter, demonstrated that stimulus imagery has a greater effect on the learning of paired associates than does response imagery. Furthermore, nouns are generally better at eliciting images than are adjectives. Paivio's finding that nouns served as more effective stimuli than

adjectives is predictable on this basis alone. Had he controlled for imagery and presented high imagery adjectives with low imagery nouns, his findings might have been very different. (A fuller discussion of imagery appears later in this chapter.)

All these objections dictate caution in the acceptance of the conceptual peg hypothesis. Furthermore, because there are so many problems associated with the study, Paivio's claim, that a well-established syntactic pattern is being overridden by a strong semantic factor becomes spurious.

The purpose of the present experiment was to eliminate some of the inconsistencies apparent in the results reported by Epstein (1961), Glanzer (1962), Wind and Davidson (1969), Baker and McCarthy (1975), and Paivio (1963). By carefully controlling for the word attributes: meaningfulness, abstractness, frequency, and imagery, and at the same time presenting the pairs in a context designed to maximize syntactic facilitation, it was hoped that more precise information about the nature (indeed about the existence) of constituent bonding could be obtained.

The Semantic Parameters

Meaningfulness. Psychologists have developed essentially three procedures for measuring meaningfulness. The first was designed by Glaze in 1928. Using approximately 2,000 CVC nonsense syllables, Glaze asked 15 subjects to indicate which nonsense items had an association for them even if they could not articulate what that association was. The association value, or "meaningfulness" measure for each syllable was simply the proportion of the subjects who reported that the syllable meant something. Hull (1933) and Krueger (1934) followed Glaze's procedure using a larger number of CVC's and subjects, and Witmer (1935) measured the association value of 4,500 CCC trigrams. The production measure used by all of these investigators is called the association value (av).

The second procedure, which is more widely used today, was developed by Noble (1952; Noble and Parker, 1960). For 96 disyllables, a mixture of real words and paralogs, he obtained continued associations and then determined the mean number of responses given to each stimulus in a 60 second time period. The maximum number of responses possible was 30, the number of blank spaces provided for each 60 second task. This figure, according to Noble, provides a meaningfulness value. Cofer (1971, p.852) reported that Noble obtained, for example, an m value of .99 for gojey and

9.61 for kitchen.

The third procedure for measuring meaningfulness is a rating technique such as the one used by Noble, Stockwell, and Pryer (1957). These investigators asked their subjects to rate, on a 5-point scale, each of 100 nonsense syllables in terms of "the number of things or ideas it made them think of" (Cofer, 1971, p. 852). The investigators then calculated measures of association value and scaled meaningfulness. Scaled m' (meaningfulness) values were obtained by converting weighted frequencies for the five-point scale into deviates of the normal curve. The same methodology was used by Noble (1961) and similar methodology by Archer (1960). All the studies mentioned here, with the exception of Noble (1952), used only nonsense items. Although Winnick and Kressel (1965) obtained m values for 32 nouns, the first attempt at obtaining meaningfulness values for a large number of real words was Paivio's (Paivio, Yuille, and Madigan, 1968) measurement of m values for 925 English nouns. Paivio followed Noble's (1952) production method, but gave subjects only 30 seconds to associate to each word. The instructions and the presentation sheets were carefully prepared to reduce the possibility of continuous (rather than continued) association. (Recently Paivio has collected normative data, as yet unpublished, on 2800 words from various syntactic categories. These data include m ratings, imagery and

abstractness values, and a rating of familiarity.)

In continuous, or chained, association, each response serves as a stimulus for the next response. For example, if the word *cat* is given as a stimulus, the response might be *dog* which would then serve as a stimulus for the next associated response, *bone*, etc. Clearly, at the end of 60 seconds the responses would have little to do with *cat*. To compute *m* values, it is necessary to obtain continued associations, i.e., the subject must associate to the same stimulus repeatedly.

The three measures of meaningfulness: *av*, *m*, and *m'*, have been reported to correlate rather highly. The majority of investigators (e.g., Goss and Nodine, 1965; Noble, Stockwell and Pryer, 1957; Underwood and Schulz, 1960) report correlations of between .70 and .90. These investigators believe that the correlations are sufficiently high, even considering the unknown variation in subject samples, experimental conditions, procedure, etc., to indicate that the three measures are all indicative of the same underlying variable (Cofer, 1971, p. 853). Paivio's (1968) factor analysis, on the other hand, failed to reveal a single dimension corresponding to these indices of meaningfulness.

Because the Paivio research, as well as most related research, obtained *m* values for nouns only, very little

comparable information is available for adjectives and verbs. (One exception is the Haagen (1949) association value ratings for adjectives. These, however, are comparative ratings and as such would not be useful for the purposes of this study.) The meaningfulness values for the present study were obtained following Paivio's procedure.

For the purposes of this study, the major question regarding av, m, and m' is how they are related to verbal learning, specifically paired-associate learning. Association value (av) and meaningfulness (m) have been more extensively studied than scaled meaningfulness (m'). Research has examined the effect of these variables on the stimulus, the response, and both the stimulus and the response (Underwood and Schulz, 1960; Goss, 1963; Noble, 1963; and Goss and Nodine, 1965). According to Cofer (1971), the substantial amount of work that has been done suggests the following generalizations on the relationship of meaningfulness measures and paired-associate learning:

> (1) Pairs with an average m or av that is higher than those of other pairs are learned more rapidly than pairs with lower m or av values;
>
> (2) Learning is a function of stimulus m or av, being more rapid for higher values of these measures;
>
> (3) Learning is a function of response m or av, again being more rapid for higher values of these measures;
>
> (4) Learning seems to be more influenced by response m or av than by stimulus m or av (p. 856).

Although there are some exceptions, claims Cofer, these generalizations seem to hold for CVC's, dissyllables, and words. These generalizations are not, however, universally accepted. The fourth point, in particular has been questioned by many researchers. Paivio (1963) for example, reported results indicating that the effect of stimulus m was greater than the effect of response m. Mandler and Huttenlocher (1956) and Epstein (1962) also report the greater effect of stimulus m. In a later study, Paivio, Yuille, and Smythe reported results suggesting that m "was equally effective on both sides of pairs when varied among abstract nouns, but had essentially no effect when the nouns were concrete" (1966, p. 362). These conflicting results regarding the effect of m may be due to a confounding of m with imagery or abstractness. If there is confounding, one possible source is the type of association that is made.

Ervin-Tripp (1967) found that in word association tasks, very young children nearly always make syntagmatic responses while adults are far more likely to make paradigmatic responses. To the stimulus dog, for example, a young child would probably respond with bark or bite. The adult response to the same stimulus would likely be bone or cat, almost always preserving the same syntactic category as the stimulus. It is entirely possible that this phenomenon is not constant across syntactic categories or even across semantic classes, e.g., concrete/abstract or high/low

imagery. If this is the case, then concrete nouns might have an advantage over abstract nouns in the traditional measure of association value since they are capable of eliciting a large number of both syntagmatic and paradigmatic responses. Abstract nouns, on the other hand, would have a severely limited number of verb responses available to them (honesty is, but not honesty sits, or honesty goes). It might be argued that a response is a response and that whether it is paradigmatic or syntagmatic is of no importance. It is important, though, in trying to establish the relationships among abstractness/concreteness, imagery, and the various measures of meaningfulness. Certainly this is an area bearing further investigation. Until control of other variables is better, any firm conclusion as to the greater effect of stimulus or response m is impossible.

Abstractness. Although researchers have agreed on the necessity of obtaining normative data on the abstractness/concreteness dimension of words, there has been little agreement as to either the definition of abstractness or the preferred procedure for obtaining the information. Darley, Sherman, and Siegel (1959) defined abstract nouns as "general, generic, not specific," but also mentioned lack of sensory experience in their instructions to their subjects. Gorman (1961) referred only to this latter

definition, instructing her two judges to consider as concrete only those nouns "whose reference to objects, to material, to sources of sensation is relatively direct" (p. 24). She obtained abstractness/concreteness judgments on 1791 nouns from two judges working with dictionaries. The judges were instructed to distinguish the two classes, i.e., to make their ratings on a 2-point scale, and were given very explicit instructions for rating "borderline" cases. Furthermore, they were instructed to subscript each judgment ("C" or "A") with an m if they believed the word to have meaning associated with the category not chosen, and ? if they were uncertain about the category assignment. The judges agreed on the categories of 95.6% (1713) of the nouns, and on both category and subscript in 80.5% (1444) of the cases. Gorman reported the 1061 nouns classified without qualification by both judges and considered them as constituting "an operational definition specifying two points on a psychological scale of abstractness" (Gorman, 1961, p. 24). Even though most investigators have found significantly high correlations between imagery and abstractness, it should be noted that Gorman confounded abstractness and imagery, as her instructions, supposedly defining abstractness for nouns, in fact related to imagery.

While it may be the case that psychological abstractness/concreteness is dichotomous, there is some evidence

that this is not always true. Gorman realized that something more sensitive than a 2-point scale was needed when she permitted her judges to subscript their category assignments. So did Spreen and Schulz (1966) when they required their subjects to rate abstractness/concreteness on a 7-point scale for 329 nouns. They defined abstractness in terms of sensory experience, and instructed their 58 subjects accordingly. Paivio et al. (1968) used essentially the same procedure for obtaining abstract ratings on 925 nouns from 28 subjects. Both the Spreen and Schulz and the Paivio results suggest that the abstractness/concreteness dimension is not strictly dichotomous. Both studies reported a great many mean values in the middle range of the scale, i.e., 3-5. Since Spreen and Schulz did not report standard deviations, it is impossible to know how much agreement was exhibited by subjects on the words with mean abstract values in the 3-5 range. Paivio, however, did report standard deviations, which provided some interesting information. For the high (6-7) and the low (1-2) mean scores the standard deviations were necessarily low, indicating relatively high agreement among subjects as to the abstractness or concreteness of a particular word. For the mid-scale mean scores, however, the standard deviations were consistently high, indicating little consensus among subjects.

Baker and Hull (1976) used the same experimental

procedure as Paivio et al. (1968) in obtaining abstractness judgments on 100 nouns chosen from Paivio's list as well as 100 adjectives and 100 verbs. They used a 7-point scale with the ends labelled "concrete" (number 1) and "abstract" (number 7), a reversal of the labelling used by earlier investigators. This reversal was motivated, in part, by the observation that while concreteness seems a fixed, absolute property, there seem to be degrees of abstractness possible. In other words, a word is either concrete or not, but if not, may fall anywhere along a continuum of abstractness. It would not, therefore, be possible to say that lemon is either more or less concrete than tree, but it would be possible to rate virtue as more abstract than formation.

In replicating the nouns from the Paivio study, Baker and Hull were motivated by four aspects of the Paivio study: (1) the high standard deviations of the middle values; (2) the fact that nouns were not controlled for syntactic ambiguity, e.g., answer, bowl, charter, and assault are commonly used as verbs in English; (3) there seemed to be some interaction between humanness and concreteness in the ratings, e.g., busybody, candidate, and evangelist received mean ratings of 5.00, 5.53, and 5.95 respectively while cat, ambulance, apple, and arrow all received mean ratings of 7.00; (4) the scaling procedure was done only on nouns. Baker and Hull attempted to reduce syntactic ambiguity by choosing nouns that do not function as either

verbs or adjectives, by using the infinitive marker to in the presentation of the verbs, e.g., to write, to confer.

The effects of abstractness on paired-associate learning have been extremely difficult to determine because of the confounding of abstractness with other factors such as imagery. Richardson (1975a) pointed out that there have been few attempts made to separate imagery and abstractness, and that this lack of interest has been justified on somewhat dubious grounds. In the first place, "there is the general theoretical attitude adopted by Paivio (1969) that the two variables are alternative measures of the same underlying process" (Richardson, 1975a, p. 215). In the second place, researchers (e.g., Paivio, 1969) have emphasized the high correlation between the two measures, generally considering separating them as impractical. Finally, items which seem to differentiate the two variables have been dismissed as peculiar (Paivio, 1969; Yuille, 1968). Richardson maintained that the justifications were fallacious. "Theoretical positions always require empirical justification, and Paivio's position is not supported by the results, since the correlation between I and C is never perfect, usually being around 0.8" (Richardson, 1975a, p. 215). The less than perfect correlation cannot be blamed on unreliable measurement of the variables since the measures "show considerable stability across different groups of subjects" (Richardson, 1975a, p. 215). But Richardson's

objection seems naive in that it is based on the assumption of an unattenuated correlation. If, as he suggests, the measurements show considerable stability across groups of subjects and are thus highly reliable, then the attenuated, or corrected, correlation might be quite high. If, for example, the reliabilities of imagery and concreteness were both .8, then the corrected correlation between them would be 1.0.

Which view is adopted, Paivio's or Richardson's, makes no difference to the question at hand. In the first case, if the variables are hopelessly confounded, there is little that can be said with confidence about the effects of a single variable on PA learning. But even if it is the case that abstractness is independent of imagery, there has been little research which investigates its effect on PA learning. Richardson (1975a,b) provided some valuable clues as to the effects of abstractness on free recall, but there seems to be little research yet on PA learning.

One experiment investigating the effect of stimulus abstractness on associational fluency, i.e., the ease with which an item elicits verbal associations, though not directly related to PA learning, bears mention because it sheds light on the properties of verbal stimuli. Lambert (1955) investigated the hypothesis that concrete nouns will elicit more associational responses than abstract nouns or adjectives. Lambert effectively conducted two simultaneous

experiments, one in English and one in French. His results in both languages supported his experimental hypothesis: concrete nouns were more effective at eliciting associational responses than were abstract nouns or the adjectives, but there was no difference between the abstract nouns and adjectives. Lambert's findings are important because they increase "the precision of selection of verbal stimuli for use in such areas as verbal learning. . . . For example, concrete words might prove to be more quickly memorized and longer retained than abstract ones. . . " (p. 106). Should this be the case, the implications for PA learning are obvious.

Word Frequency. Although a number of word counts have been published, only two are comprehensive enough to be widely used for research purposes. One is the Thorndike-Lorge (1944) count which combined data collected in a number of word counts over several years' time by two authors. The data, a total of 4.5 million words, came from magazines, e.g., The Saturday Evening Post, True Story, and Readers' Digest, and from children's and juveniles' books, e.g., Black Beauty and Little Women. Of the 30,000 words listed in the Thorndike-Lorge book, 19,000 occurred at least one time per million. Words occurring more than 100 times per million were designated AA, and those occurring 50-100 times per million were designated A. No precise frequencies were

given for these words.

The Thorndike-Lorge book is considered to be useful and is widely employed by researchers. It has, however, several shortcomings. The major one is its age. More than three decades have passed since its publication and more than five since some of the counts were made. The language has changed since then, a fact which probably would not alter significantly the tallies for the very high frequency words, but which would change considerably the tallies for the low frequency words. A second flaw in the book is the fact that precise frequencies for the high-frequency words are not given. These data are necessary for many researchers whose work depends on precise information.

These shortcomings are overcome in the Carroll, Davies, and Richman Word Frequency Book (1971). The data in this book were collected from over 1000 different publications representing over five million words of running text published since November of 1969. The book lists 85,000 different words, and actual raw word counts, as well as ranking information, are given for each.

All of Paivio's research and most other research involving word frequency as a variable used the Thorndike-Lorge counts. The more recent Carroll counts were used in the present study.

The significance of variation in word frequency to

paired-associate learning has not been widely studied. Briefly, the available literature suggests that:

> (a) acquisition was faster . . . when the response terms were high-frequency words than when they were of moderate frequency;
>
> (b) the slowest acquisition occurred when low-frequency response words were to be learned;
>
> (c). . . when the word frequency was varied on the stimulus side, learning rate was best for words of moderate frequency. (Cofer, 1971, p. 896)

It is important to note that these conclusions were reached on the basis of studies using word-word pairs, and not on word-nonsense syllable pairs. There is, however, no reason to suspect that different results would be obtained from word-nonsense syllable pairs.

Imagery. There has never been much argument among psychologists that the ability to form mental images is a powerful factor in verbal learning. Paivio traces the history of imagery in psychology as follows:

> As every psychologist knows, imagery once played a prominent role in the interpretation of (associative meaning, mediation, and memory). It was widely regarded as the mental representative of meaning - or of concrete meaning at

> least. William James, for example, suggested that the static meaning of concrete words 'consists of sensory images awakened' (1890, p. 265). As manifested in the 'wax tablet' model of memory, imagery was the prototype of stimulus associative imagery, it was assumed to play a mediational role in mnemonic techniques which originated long ago as a practical art (see Yates, 1966). (Paivio, 1969, p. 241)

Despite strong criticism directed against such views, they have continued to gain acceptance as being intuitively plausible.

According to Ellis (1972), mental imagery has been studied in two general ways. One way has been to instruct subjects to construct mental images that will relate the two words to be associated. The second way (and the one that was employed in the present study) has been to vary the imagery values of the verbal materials, real or nonsense words, in a learning experiment.

The traditional procedure for obtaining imagery values is to ask subjects to rate words or nonsense syllables according to the ease with which they generate mental images. This is the procedure used by Paivio (Paivio and Steeves, 1967; Paivio, Yuille, and Madigan, 1968) and by Baker and Hull (1976) in obtaining the imagery values for the present study.

That imagery facilitates PA learning is a well-established fact in verbal learning research. Paivio has

investigated many aspects of imagery, but perhaps the most germane are the effects of imagery instructions and his notion of the conceptual peg.

The effect of imagery instructions has been clearly established (Paivio, 1969; Paivio and Yuille, 1967; Paivio and Foth, 1970), and is easily understood. In simplest terms, the assumption is that memory for concrete objects is somehow superior to memory for words, and that instructing subjects to form mental images (assuming that they follow the instructions) will improve PA learning. Paivio and Foth (1970) offered some experimental support for this assumption. They required their subjects to generate mediators for noun pairs. Half the mediators were to be verbal and half imaginal. Their results showed that imagery produced better recall than did verbal mediation for concrete nouns. For abstract nouns the reverse situation occurred. The problem with this study is, of course, that subjects may very well retain verbal mediators even with imaginal ones. Instructed to form an imaginal mediator for the noun cow, does the subject only form a mental image of a cow, or does he both form the image and retain the word? If he does use a verbal mediator, then a study of this type is of only slight value.

Most researchers to date have found it to be the case that stimulus imagery has a greater effect on PA learning than does response imagery. Thus in a real-word PA task,

the imageability of the stimulus word exerts more influence on learnability than does the imageability of the response term. The conceptual peg hypothesis is certainly consistent with these findings (see Paivio, 1963; 1969). "The argument is that the stimulus member of a pair serves as a 'conceptual peg' . . . to which its associate is hooked during learning trials when the stimulus member is presented alone" (Paivio, 1969, p. 244). Assuming that imagery serves as a mediator, the ease of learning will depend, partly, on the capacity of the words in the pair to evoke images. The image-arousing value of the stimulus word should be particularly important because it must function as the cue that reinstates the compound image formed during the learning process (Paivio, 1969).

Two experiments done by Butter and Palermo (1970) attempted to investigate the effect of concreteness on paired-associate recall. They measured both immediate recall and recall 48 hours after presentation. In both cases, the results indicated significant effects of stimulus concreteness and response concreteness, with stimulus concreteness exerting more influence. The investigators made no attempt to control for imagery or to consider the effects of imagery. They merely assumed Paivio's hypothesis that imagery and concreteness measure the same variable and interpreted their results as support for Paivio's (1969) theory that concrete nouns are superior to abstract nouns in

facilitating recall because of their image-arousing capacity. But, of course, their conclusion is spurious. The same argument would hold for any property of the stimulus since stimulus imagery, like the stimulus, is always present as a given and is inseparable from it. All Butter and Palermo really demonstrated was the direction of the bonding, i.e., from the stimulus to the response rather than from the response to the stimulus. If high imagery stimuli are better, then any property co-varying with imagery is also a candidate for explanation.

These findings with regard to stimulus imageability are important because they contrast with what most reseachers have found to be the effect of meaningfulness. While the issue is by no means settled, as noted above, learning is generally thought to be more influenced by response meaningfulness than by stimulus meaningfulness. The differing effects of the two parameters on PA learning suggests that they are not highly correlated, and indeed, the research that has been done in the area, e.g., Paivio, 1968 and Paivio and Yuille, 1967, has failed to establish strong correlations between I and m.

Kintsch (1972) pointed out that there may be something besides abstractness that confounds most imagery studies reported in the literature. Because imagery and concreteness are very highly correlated, most of the low imagery nouns used in experiments are abstract. Kintsch

suggested, however, that there may be something more than imagery which makes low imagery nouns harder to learn in learning experiments than high imagery nouns. He assumes that people have as part of their long-term memories some kind of a dictionary which identifies words and defines their meaning. It has been suggested by both psychologists and linguists interested in the structure of this dictionary, that some words of a language are simple lexical items while others must be decomposed into simpler, basic units. Many abstract words, e.g., destruction, conception, wisdom, fall into the latter category while most concrete words fall into the former. In a series of three experiments, Kintsch found a significant effect on learning that was attributable to lexical complexity independent of imagery, abstractness, and word frequency.

While Kintsch has made an interesting point, he has failed to make an important distinction between linguistic and psychological complexity. His exemplars are certainly linguistically complex, but there is no reason to assume that they are complex in any processing sense for the speaker/hearer. Certain other items are linguistically complex in a different way, but are also clearly psychologically complex. Unfilled and unrealized, for example, are, from a processing point of view, psychologically more complex than destruction.

Purpose of this Study

This chapter has reviewed a series of studies which investigated constituent bonding. Stated in general terms, the problem was to discover the properties which govern the learnability, or the "bondability," of words. The research done thus far has investigated two broad areas which influence this bonding. Paivio, Butter and Palermo (1970), and Lambert (1955) investigated what might be called the semantic factors while Epstein, Glanzer, Kanungo, and Baker and McCarthy investigated the syntactic factors. It is tempting to state that the former has been the domain of the psychologists and the latter the domain of the linguists. But, in fact, the two disciplines have begun to interact, at least regarding the problem of constituent bonding, and reseachers are starting to recognize that they must account for both semantic and syntactic factors in any serious explanation of the phenomenon (cf. Baker and McCarthy, 1975).

The present study represents an attempt to integrate what has been two separate areas of investigation. There is no doubt that the syntactic properties of words influence their learnability. The experimental results, however, have not been entirely satisfactory. The Baker and McCarthy study, for example, demonstrated the existence of a phenomenon that is probably syntactic facilitation. But

there was no satisfactory explanation for the phenomenon (although several possibilities were offered). Furthermore, there was some question as to whether the phenomenon was operative for nouns.

On the other hand, it is clear that the semantic variables alone cannot account for the varying learnability of words. What is needed is integrated experimentation - that is, experiments measuring the effects of syntactic and semantic variables simultaneously. That is the intent of the present research.

CHAPTER II

PROCEDURE

To understand the relevance of the experimental procedures of this experiment to the problem of semantic organization, it is neccessary to review the processes involved in paired-associate (PA) learning. PA learning involves the presentation of a list of paired items. During successive trials the subject is given one member of the pair and is expected to recall the other. It is generally recognized that there are three stages in PA learning: learning the stimulus item, learning the response item, and associating or bonding them together. In order to do this last stage, the subject must form some kind of bond which uniquely associates the stimulus and the response. He must, in other words, construct a single unit from two previously unconnected parts. The subject, in searching the semantic network for a basis on which to bond the two items, can be forced to play with a stacked deck. Certain routes can be blocked or partially blocked. For example, most phonological similarity, in the present experiment, was eliminated between members of a pair. Other routes were encouraged. The use of syntactic information was maximized in some instances by providing very obvious syntactic frames in which to learn the pairs.

The paradigm for the present experiment was the traditional PA procedure used by Glanzer, Kanungo, Baker and McCarthy, etc. One important change was to present all words in both positions, preceding and following the nonsense syllables. When half of a list of words consistently precedes the nonsense words and half follows (e.g., Baker and McCarthy, Table 1), there is no assurance that syntactic facilitation alone is responsible for the learning differences. It may be that the words in one list share some common property that affects learnability, or that the difference is due to the unique semantic properties of the individual words.

One solution is to divide the subjects instead of the word lists. This was the procedure followed in the pilot study to the present one. Half the subjects in the pilot study received a given list of paired associates in the real word-CVC order. The other half received the same list of paired associates, but in the CVC-real word order. The results, however, indicated no significant effect due to position as was predicted. Despite the fact that the ADJ-CVC position would be the expected preference (i.e., more easily learned) for English speakers, the data for 24 subjects showed no such preference. For this reason it was decided to adopt a different procedure, namely to divide the lists as described below.

Materials

Following Baker and McCarthy's change from the Glanzer and from the Kanungo studies, only real words were used as stimuli. There are several reasons for this alteration. In the first place, their syntactic information is consistently presented from trial to trial. It is easier, furthermore, to control equality of difficulty for responses if they are nonsense syllables than if they are real words. Finally, syntactic bonding can extend from the real words to include the nonsense forms, but cannot extend from the nonsense forms which have no syntactic information to supply. The fact that the Baker and McCarthy study was able to demonstrate learning differences that might be attributable to syntactic facilitation more clearly than the previous studies demonstrates that the motivation for this mode of presentation is reasonable.

Three lists of 12 real word-CVC pairs were constructed, one each for nouns, adjectives and verbs. The words were selected from the 300 words (100 each of nouns, adjectives, and verbs) for which Baker and Hull (1976) collected imagery and abstractness ratings. The real words were chosen to be non-ambiguous members of the grammatical category they were intended to represent, e.g., syntactically ambiguous items such as chair which could function as a noun or as a verb, were eliminated. Also eliminated were verbs whose past

participial forms are used commonly as noun modifiers, e.g., tired, and potentially homophonous items, e.g., wee. Words were chosen to represent the polar values of the imagery scale (Baker and Hull, 1976) and, because the correlation with concreteness is very high, the abstractness-concreteness scale as well. Half the items in each list were very high in imagery value (and low in abstractness) and half were very low (and highly abstract). All the verbs chosen were transitive, regular, and two or three syllables in length. They were presented in their past tense forms. Because of the I value and frequency restrictions placed on the adjectives, it was impossible to have all of them be of the same length. Consequently, the length ranged from one to five syllables. None of the nouns exceeded four syllables; most were two syllables. (The full list is given in Table 1 in Chapter III.)

The nonsense syllables were consonant-vowel-consonant (CVC) types with a rated association value of 27% or less on the Glaze lists (Hilgard, 1951, pp. 543-544). The CVC responses were selected and paired randomly with the real words such that each subject received a unique pairing. Within each list no two CVCs began or ended with the same letter. (The full list of CVC's is given in the Appendix.)

Methods

Each of the three lists (one list each of nouns, verbs and adjectives) was prepared for presentation as follows: A master list was prepared for each syntactic category. Each list contained 12 words paired randomly with 12 nonsense syllables. For each presentation there were two options for positioning of the real word relative to the CVC. If the experimenter specified "R" order, the first six words of the master list were presented to the right of the CVC, and the second six to the left. Specifying "L" meant that the first six items of the master list were presented to the left of the CVC and the second six to the right.

The pairing of the CVCs with the real words was cyclical such that the first subject received the first real word paired with the first CVC, the second real word with the second CVC, etc. The second subject received the first real word paired with the second CVC, the second real word with the third CVC, and the twelfth real word with the first CVC, etc. until the twelfth subject received the first real word with the twelfth CVC. The thirteenth subject saw the same pairings as the first subject, as did the twenty- fifth and thirty-seventh subjects. For each subject, then, the experimenter specified presentation order, "R" or "L" for each list and pair order, 1-12, as well as the order in which the lists were to be presented (Noun, Verb, Adjective; Verb, Adjective, Noun, etc.). Order of pairs within lists

was randomized before each presentation.

The presentations to subjects were made on the CRT screen of a PDP 12 computer for learning by the "prompting and anticipation" technique. Each item was set up for a three-stage display: the preparatory signal consisting of two asterisks, the stimulus (real word) presentation with asterisks indicating the position to be occupied by the response, and then the stimulus and response pair. Subjects were required to read the items aloud, left to right, as they appeared, and they were required to try to anticipate the response at the second stage. Within the lists the pairs were re-ordered for each presentation to eliminate the possibility of serial learning. Prior to the experiment proper, each subject learned a practice list in order to familiarize him with the technique to be used.

Subjects had ten seconds during which to respond before the correct response was given automatically by the computer. If a response was given before the end of the ten second period, the experimenter instructed the computer, via a remote control switch, that a response had been made and indicated whether it was correct or not. Learning on a list was continued until the subject could anticipate 24 consecutive pairs correctly.

The responses were coded and each subject assigned a percent correct for each word determined by the following

formula:

$$\frac{NC}{30} \pm \frac{(30-NP)}{30} \times 100 = \% \text{ correct}$$

(NC=Number Correct; NP=Number of presentations)

No subject required more than 29 trials to reach criterion on any list. Assuming that a subject would continue giving correct responses if testing continued beyond that, the percent correct scores would be achieved if all subjects had been given a standard 30 trials. In order to assure proper distributional properties for subsequent analyses of variance, the percent correct scores were subjected to an arc sine transformation. It is these latter values which appear in the Results section of this report, and these are hereafter referred to as the learnability scores for each item. Response latencies were recorded by the computer in centi-seconds, but it is the natural logarithms of the actual times that are reported in the next chapter.

Subjects

A total of 48 subjects drawn from introductory courses in linguistics and psychology were used. They ranged in age from 20 to 45 years with the average age being 24. All were native speakers of English with normal vision and hearing, and none had previously participated in a paired-associate

learning experiment. There were thirty females and eighteen males.

CHAPTER III

RESULTS

The primary purpose of this study was to explore some specific properties, both syntactic and semantic, that influence the learnability, or more specifically, the bondability of certain words. The question comes down to this: why were some of the words the subjects learned in the lists learned more easily than others, i.e., what properties do the more easily learned words share?

The data collected presented a rather complex analytical problem. With so many possible statistical approaches to the analysis, it was difficult to determine which gave the experimenter the most accurate portrayal of what actually occurred in the experiment. The best solution seemed obviously to focus on the words, to consider learnability as measured in the experiment as a quantifiable variable of the words in the same way as frequency, imageability, and meaningfulness. This analytical strategy provided a great deal of information.

Table 1 summarizes all the information collected on each word before and during the experiment described. Imagery, abstractness, meaningfulness, and frequency scores are given in the first four columns followed by the mean

TABLE 1

Words	Imagery	Abstractness	Meaningfulness	Frequency	Learning Scores			Response Latencies		
					Left	Right	Pooled	Left	Right	Pooled
Forest	6.71	1.73	8.23	2.34	71.15	71.14	71.14	3.32	3.39	3.35
Orchestra	6.55	1.81	7.36	1.77	72.52	70.70	71.61	3.20	3.32	3.26
Instrument	6.50	1.74	7.48	0.91	72.96	74.30	73.63	3.21	3.18	3.19
Virtue	2.32	6.03	4.68	0.84	69.12	68.90	69.01	3.43	3.36	3.40
Heredity	1.93	4.87	6.02	0.67	68.18	70.01	69.10	3.69	3.56	3.63
Formation	2.61	4.77	5.20	1.24	68.22	68.85	68.53	3.50	3.41	3.45
Capacity	2.32	5.42	5.07	1.15	68.55	69.54	69.04	3.57	3.48	3.52
Immunity	2.13	5.13	5.55	0.29	65.61	68.15	66.88	3.62	3.55	3.58
Opinion	2.23	5.87	5.18	1.70	68.35	67.10	67.73	3.37	3.52	3.44
Lemon	6.61	1.52	7.84	1.09	72.15	72.49	72.32	3.32	3.11	3.21
Hotel	6.55	1.71	6.91	1.56	72.72	72.84	72.78	3.25	3.23	3.24
Library	6.68	1.68	6.98	1.85	71.42	70.91	71.16	3.21	3.32	3.26
Tall	6.55	2.19	6.39	2.22	68.75	71.16	69.96	3.33	3.28	3.31
Filthy	6.23	2.29	6.22	0.37	70.19	67.35	68.77	3.42	3.62	3.52
Peppery	6.13	2.52	6.34	-0.71	72.69	71.57	72.13	3.37	3.28	3.32
Coherent	2.97	4.45	4.75	-0.11	67.34	67.84	67.59	3.48	3.42	3.45
Simple	3.19	4.16	5.91	2.89	69.16	69.10	69.13	3.43	3.36	3.39
Bad	3.26	4.90	7.20	2.12	70.20	69.14	69.67	3.36	3.38	3.37
Reliable	3.45	4.61	5.18	0.85	66.25	65.23	65.74	3.52	3.66	3.59
Imaginary	3.42	5.32	5.93	1.26	68.54	66.83	67.69	3.45	3.58	3.51
Fickle	3.16	4.61	4.84	0.07	69.14	66.51	67.83	3.48	3.55	3.52
Leaky	6.23	2.26	6.16	0.29	69.97	67.58	68.78	3.25	3.40	3.32
Cozy	5.61	3.71	6.89	0.72	70.93	69.48	70.21	3.33	3.40	3.37
Clear	5.64	2.81	6.30	2.20	68.74	68.04	68.39	3.35	3.40	3.37
Dissected	5.73	2.17	6.43	0.29	65.73	66.96	66.35	3.59	3.44	3.52
Suffocated	5.84	2.35	5.86	0.07	64.33	66.21	65.27	3.55	3.50	3.52
Carried	6.10	2.45	6.39	2.48	66.14	65.83	65.98	3.42	3.44	3.43
Conceded	3.06	3.97	4.27	0.44	64.19	62.40	63.29	3.69	3.56	3.63
Predicted	3.16	4.29	5.27	1.29	61.97	61.83	61.90	3.60	3.66	3.63
Injured	5.68	2.64	7.50	1.23	67.75	67.29	67.52	3.34	3.36	3.35
Sustained	2.84	4.42	4.82	0.89	61.89	61.82	61.85	3.71	3.72	3.72
Relieved	3.13	5.66	5.09	2.35	66.59	65.33	65.96	3.40	3.57	3.49
Expected	3.27	5.13	4.96	2.15	63.97	64.55	64.26	3.55	3.58	3.57
Devoured	6.23	2.39	6.05	0.75	67.17	66.59	66.89	3.51	3.47	3.49
Navigated	5.48	2.23	6.61	0.44	67.99	72.18	70.09	3.36	3.24	3.30
Considered	2.81	5.06	4.36	2.17	62.49	61.60	62.04	3.65	3.70	3.67

learning scores (as defined earlier) for the word when presented on the left, the mean learning score for the word presented on the right, and the mean learning score for the word averaging over left and right positions. The final three columns give the mean response latencies (measured as described in the previous chapter) for each word presented on the left, presented on the right, and averaged over both positions.

A few points need to be made which may help to clarify the analysis to follow. First, the imagery values for the words listed here are not representative of the entire range of imagery values. Rather, they represent the high and low extremes of the seven point scale. Because the standard deviations for the middle values were quite high, it was decided that greater precision of interpretation would be obtained by excluding them from consideration. Second, the words to be used in the experiment were chosen for their imagery values (and, of course, their syntactic category), but no restrictions were placed on the abstractness values. Because imagery and abstractness are so highly correlated, however, the abstractness scores represented polar extremes as well, once again on a seven point scale.

The meaningfulness scores given here represent the mean number of associative responses given by subjects to each word during a thirty second time period as determined in the Baker and Hull (1976) study. Frequency scores are reported

as the log frequency per million based on the Carroll, Davies, and Richman (1971) counts. A word occurring less often than once per million words will, thus, have a negative frequency in this table.

Learning scores are the arc sine transformations of percent correct scores for all the subjects. Finally, it is the natural logarithms of the response latencies that are given in the last three columns. The figures given in Table 1 form the basis for the following analyses.

The Pearson Correlation Coefficients for each variable correlated with every other are given in Table 2. These figures provide a great deal of information and must be examined systematically.

The correlation of imagery with abstractness is -.95, significantly higher than the .83 reported by Paivio et al. (1968) for 925 nouns. Furthermore, correlations computed on the present data between imagery and abstractness within each syntactic category showed a high degree of consistency, -.98 for nouns, -.94 for adjectives, and -.93 for verbs. These figures provide even stronger support than Paivio's data for the conclusion that the two scaling measures "appear to be defining a common dimension of word meaning" and that "for research purposes, either scale can be used for item selection" (p.7). (It should be recalled that the negative correlation is attributable to the reversing of the

TABLE 2

PEARSON CORRELATION COEFFICIENTS

	Abstractness	Meaningfulness	Frequency	Learning Scores			Response Latencies		
				Left	Right	Pooled	Left	Right	Pooled
Imagery	-.95	.77	.01	.52	.49	.52	-.68	-.61	-.69
Abstractness		-.75	.07	-.45	-.49	-.48	.56	.61	.62
Meaningfulness			.17	.68	.70	.71	-.71	-.69	-.75
Frequency				-.00	-.03	-.02	-.24	-.04	-.15
Learning Scores (L)					.88	.97	-.82	-.77	-.85
Learning Scores (R)						.97	-.70	-.87	-.84
Learning Scores (Pooled)							-.79	-.85	-.87
Response Latency (L)								.75	.93
Response Latency (R)									.94

concreteness/abstractness scale in the Baker and Hull study.)

The correlation between imagery and meaningfulness across all syntactic categories is .77, very near the .72 found by Paivio et. al. (1968). Those researchers suggested that abstract words derive their meaning largely from intraverbal experience (Paivio, 1966), so that the less than perfect correlation might be due, partly, to the fact that certain nouns with highly emotional or evaluative connotations were high in rated meaningfulness but low in both imagery and concreteness. A somewhat different light is shed on this explanation by the correlation coefficients of imagery and meaningfulness considered within each syntactic category in the present study.

For nouns, the correlation between I and m is .91, significantly higher than Paivio's .72. Since the list contains no highly emotive or affective terms (see Table 1 or Appendix A), this high correlation might be consonant with Paivio's data if the latter were re-examined after controlling for the affective values. Unfortunately, Paivio did not report a correlation coefficient for the altered data, so direct comparisons with the present study are impossible (Paivio et al., 1968). Similarly, the correlation of imagery and meaningfulness for verbs is .86, and the verb list contained no items which might be judged higly emotional or evaluative.

For adjectives, however, the correlation between I and m was only .51, suggesting that the relationship between the two properties is not constant across all syntactic categories. At first glance, Paivio's explanation seems of little help here since the items are not particularly emotive or affective, with one exception. Bad has the highest m value in the list, 7.2, (the average of the other 11 is only 5.9), but is relatively low in imagery (3.26, see Table 1).

The aberrant behavior of bad, the only item in the list with emotional and evaluative connotations and with a highly inflated m value compared to the other adjectives, goes a long way toward explaining the low correlation of I and m within adjectives. The correlation of I and m with bad removed from consideration is .79, roughly the same as for the verbs, and approximately the same as Paivio's .72 which still included the emotive terms. This last point is important. It casts doubt on Paivio's attempt to attribute the less than perfect correlation to the emotive terms in the list and imply, thereby, that for non-emotive terms the relationship between imagery and meaningfulness might be stronger than his .72 correlation coefficient suggests. The correlations between imagery and meaningfulness in the present data for verbs and adjectives without emotive terms are not significantly different from Paivio's. For nouns, however, the difference is significant. Clearly there is

something in addition to the affective nature of words differentiating imagery and meaningfulness. While further research might show that this property serves to differentiate the two variables in nouns, the data examined here, when compared with Paivio's (Paivio et al., 1968) require that additional explanations be sought.

The correlation between abstractness and meaningfulness is -.75, almost identical to the correlation between imagery and meaningfulness. However, two points need to be made. First, Paivio et al. (1968) found meaningfulness more closely related to imagery (r=.72) than to concreteness (r=.56). They attributed the relatively low correlation of meaningfulness and concreteness to the asymmetry in the relationship in words with emotional connotations, i.e., the same explanation invoked for the correlation between imagery and meaningfulness. Second, further analysis in the present study revealed that the correlations of the two variables within syntactic categories followed the same pattern as the correlations between imagery and meaningfulness. Specifically, the correlation between abstractness and meaningfulness within syntactic categories was -.94 for nouns, -.39 for adjectives, and -.79 for verbs. Removing bad from consideration raised the correlation coefficient for the adjectives to -.61, closer to the other two syntactic categories, but enough lower to suggest that the relationship between meaningfulness and abstractness might

not be consistent across syntactic categories. But more important is the observation that with regard to the semantic properties under consideration here, the interrelationships appear not to be constant even within the category, adjective. Although the evidence reported here is rather too tentative to be considered conclusive, it seems that emotional and evaluative terms do behave differently under experimental conditions than do other terms. Paivio's data for nouns support this intuitively obvious but mostly unarticulated position, and the data on adjectives in the present study offer some support. Certainly further experimentation is warranted. The inconsistency of the interrelationships among the syntactic categories and within the adjectives is a subject which will be explored later in this chapter and in the discussion which follows.

The correlations between frequency and all other factors is non-significant. The non-significant correlation coefficients might be interpreted as evidence that, for purposes of psycholinguistic experimentation, the traditionally used word counts as measures of frequency are inadequate. Frequency counts are based on several types of written material, but it is written and thus may not reflect accurately the frequency of occurrence of words in the language where language is understood to include speech as well as writing. The relative formality of written prose as well as the limitations placed on subject matter - some

matters, for example, are too trivial to demand much written comment but may warrant considerable oral discussion - may invalidate the frequency measurements in a study of this kind.

Two additional points regarding frequency were made by Janet Taylor (1958) who pointed out that there are two highly correlated aspects of frequency. One is the fact that "words vary with respect to the sheer frequency of exposure to the letter pattern that an individual has experienced in the past," and the second is that "such words may be assumed to vary in meaningfulness, the more familiar, frequently used words possessing more meaning for the individual" (p.329). Taylor's points are important. Both suggest that a measurement of frequency based on written text may have little to do with actual frequency if that is meant to be a measure of familiarity to the language user. Factors such as recurrence of grapheme patterns and the users own linguistic experience very likely influence actual frequency or familiarity.

Paivio recognized that the traditional concept of frequency might be problematic in experimental situations, and has devised, as an alternative measure, a direct familiarity rating. Using methodology similar to that of the imageability and abstractness measures, he simply asked subjects to rate words on the basis of how "familiar" they were (Paivio, 1977, unpublished). It is not unlikely that

the very high frequency words, such as those used by Baker and McCarthy (1975), would likely be immune from such criticism, probably occurring with a high frequency in the spoken language as well as the written, while many of the words in the present study would be subject to this criticism as they represented a wide frequency range. In other words, it is more likely that the frequency measurements for Baker and McCarthy's words reflect accurate frequency (familiarity) than do the words in the present study. This in itself might help to explain the differences in the results between the two studies (to be discussed more fully below).

The correlation of meaningfulness with learnability is .71. This is the highest of the correlations between learning scores and the semantic properties, suggesting that meaningfulness was the single best predictor of learnability. This is in marked contrast to a number of earlier studies (see references in Chapter One) which have considered imagery to be the best indicator. This finding motivated further analysis, to be discussed shortly. It should be noted that the correlations of both position scores (left and right) with the semantic factors are virtually identical to the correlations of learnability averaged over the two positions with the semantic factors, a fact which justifies the consideration of only the averaged figures in subsequent analysis and discussion.

The correlation of meaningfulness and learnability within syntactic categories follows the pattern established above for abstractness and meaningfulness, and for imagery and meaningfulness. The correlation of meaningfulness and learnability for nouns is .81, for adjectives, .44, and for verbs, .79. Once again, leaving out bad raised the correlation coefficient, this time from .44 to .69. Removing bad makes the adjectives a more homogeneous group of words, i.e., like the nouns and verbs, free of items with emotional and evaluative connotations. It is tempting to claim at this point that the semantic properties under consideration have roughly the same effect on learnability and are related in the same way across all three syntactic categories as long as the words are not marked for emotionality.

Still, the fact that all the correlations that have been discussed are lower for adjectives than for nouns or verbs cannot be ignored. It suggests that other factors, possibly syntactic category, influence adjectives differently than nouns and verbs.

The correlation of imagery with learnability is .52 across all syntactic categories, and the correlations within syntactic categories are .86 for nouns, .40 for adjectives, and .77 for verbs. Removing the aberrant bad increases the figure to .64, more in line with the other two categories.

These figures are particularly interesting when compared with the correlation of meaningfulness and learnability across and within syntactic categories. Within syntactic categories, the correlations between meaningfulness and learnability and between imagery and learnability are nearly identical, but across syntactic categories, the correlations are .71 and .52 for meaningfulness and imagery respectively, a difference which is significant at the .05 level. Further discussion of the correlation between imagery and learnability will be deferred until the partial correlations are presented later in this chapter.

The correlation of response latency with the semantic properties reveals a high degree of consistency with learnability scores. (It should be noted that the correlation of response latency (left) and response latency (right) and of response latency averaged over the two positions with the semantic properties and with learnability are virtually identical, a fact which motivates using only the averaged figure in this and subsequent discussions.) The correlations between response latency and the semantic factors are roughly the same as between learnability and the semantic factors (see Table 2). These figures suggest a high degree of relatedness between the two basic response measures used in the study, and indeed, response latency and learnability correlate at -.87. This contradicts the generally held view that measures related to percent correct

and response latency are incongruous in their assessment of any particular behavior. For example, Blough and Lipsett (1971, p.747) assert that "a subject might choose an incorrect stimulus as often as a correct stimulus but take longer to choose the incorrect one. 'Percent correct' would show no discrimination, but 'latency' would." The present study showed response latency to be congruent with percent correct in measuring the variables of interest.

Table 3 gives the partial correlation co-efficients controlling for meaningfulness, and Table 4 gives the partial correlatiocn co-efficients controlling for imagery. The partial correlation co-efficient respresents a measure of that part of the correlation between two variables, e.g., imagery and learnability, that is not simply a reflection of their relationship to a third variable, in this case, meaningfulness. Specifically, the question being asked is whether imagery and learnability (or imagery and response latency) are correlated merely because of their common association with meaningfulness, or whether they have a relationship independent of meaningfulness. The partial correlation coefficient for imagery and learnability controlling for meaningfulness, as seen in Table 3, is r=-.0624, statistically non-significant.

This rather radical drop from .52 to .06 indicates that the apparent correlation between imagery and learnability is largely dependent upon, or mediated by, their mutual

TABLE 3

PARTIAL CORRELATION COEFFICIENTS

(Controlling for Meaningfulness)

	Abstractness	Frequency	Learning Scores			Response Latencies		
			Left	Right	Pooled	Left	Right	Pooled
Imagery	-.89	-.20	-.00	-.11	-.06	-.29	-.16	-.25
Abstractness		.30	.11	.08	.10	.06	.19	.14
Frequency			-.17	-.21	-.20	-.17	.12	-.03
Learning Scores (L)				.77	.94	-.65	-.56	-.70
Learning Scores (R)					.94	-.41	-.75	-.67
Learning Scores (Pooled)						-.56	-.70	-.73
Response Latency (L)							.51	.86
Response Latency (R)								.87

TABLE 4

PARTIAL CORRELATION COEFFICIENTS

(Controlling for Imagery)

	Meaningfulness	Frequency	Learning Scores			Response Latencies		
			Left	Right	Pooled	Left	Right	Pooled
Abstractness	-.07	.24	.17	-.07	.05	.40	.12	-.13
Meaningfulness		.27	.51	.58	.57	-.40	-.44	-.48
Frequency			-.01	-.03	-.02	-.32	-.04	-.20
Learning Scores (L)				.84	.96	-.75	-.66	-.79
Learning Scores (R)					.96	-.58	-.83	-.80
Learning Scores (Pooled)						-.69	-.78	-.83
Response Latency (L)							.58	.88
Response Latency (R)								.90

correlations with meaningfulness. That meaningfulness is the more basic or underlying variable is indicated by the calculation of the partial correlation for meaningfulness and learnability, controlling for imagery. That correlation, originally .71, only drops to .57 and remains clearly significant.

The pattern is similar with respect to the response latency measure. The original correlation of -.69 between imagery and response latency drops to a statistically non-significant -.25 when adjusted for meaningfulness. However, when the relation between meaningfulness and response latency, r=-.75, is adjusted for imagery, it remains significant at -.48.

Based on these figures, it seems reasonable to conclude that of the semantic factors examined, meaningfulness clearly exerted the greatest influence on the learnability of the 36 words in the experiment. If this finding holds in general, i.e., for all words, it is extremely important. It suggests that psychologists, who have focused their attention so extensively on imagery, have been looking at the "wrong" variable. Their attention should be redirected toward meaningfulness (in Noble's sense of that term, i.e., richness of association value).

Given the potential significance of the above results, and the fact that the imagery values were selected from the

extremes of a distribution, it was considered important to determine whether or not the correlations with regard to imagery were at all spurious. As a check on this, all of the correlations were recomputed from ranked data yielding Spearman coefficients. The results were uniformly the same as those produced by the Pearson coefficients. The correlations, e.g., between imagery and learnability were .52 and .50 for the two forms. For imagery and response latency, they were -.69 and -.68. A similar close match appeared for all of the variables.

The next step in the analysis was to perform an analysis of variance with the learning measure as the dependent variable. The between-subjects factors were syntactic category (3 levels, noun, adjective, verb) and imagery (2 levels, high and low). The within-subjects factor was position (2 levels, right and left of the CVC). Table 5 shows that there were significant effects due to syntactic category, imagery, and the position by syntactic category interaction. The previous analysis has indicated that the effect of imagery is via meaningfulness, and the implications for that finding have been suggested. The interesting result in this analysis was the position by syntactic category interaction. A test on simple effects revealed that it was the adjectives which accounted for the interaction effect in the ANOVA. In other words, the position of the real word relative to the CVC apparently did

TABLE 5

ANALYSIS OF VARIANCE SUMMARY TABLE
(Percent Correct)

Source	Sum of Squares	Degrees of Freedom	Mean Squares	F Ratio	Probability
Category (C)	336.422	2.	168.211	45.462	0.001**
Imagery (I)	172.453	1.	172.453	46.609	0.001**
C x I	15.961	2.	7.980	2.157	0.133
Within Subjects	111.000	30.	3.700		
Position (P)	0.445	1.	0.445	0.450	0.507
C x P	6.516	2.	3.258	3.292	0.051*
I x P	0.258	1.	0.258	0.261	0.613
C x I x P	4.594	2.	2.297	2.321	0.116
C x Within Subjects	29.688	30.	0.990		

TABLE 6

DATA ADJUSTED FOR MEANINGFULNESS

Source	Sum of Squares	Degrees of Freedom	Mean Squares	F Ratio	Probability	% Reduction
Category (C)	205.250	2.	102.630	29.490	0.001	39
Imagery (I)	20.360	1.	20.360	5.850	0.025	88
C x I	4.610	2.	2.310	0.660	N.S.	
Within Subjects	100.910	29.	3.480			10

not affect the learnability of the nouns and verbs. For adjectives, however, it was important, the preferred position being the adjective to the left of the CVC. This is predictable from English phrase structure and consistent with Baker and McCarthy's (1975) results.

The important question raised by this finding was why the adjectives demonstrated a position effect when neither of the other two categories did. It was suggested earlier that it would be difficult to predict the preferred position for nouns given that both Adjective-Noun and Subject-Verb form relatively stable units in traditional English phrase structure. It is impossible to say whether the non-significance of position for nouns was due to an equal facilitative effect, i.e., subjects formed Adjective-Noun constructions as often as they formed Subject-Verb constructions or they formed compound nouns with the CVC's in either position, or whether the semantic factors took on more salience than the syntactic in the PA learning task. Either way, the results are in accord with Baker and McCarthy (1975) who found evidence for syntactic facilitation in adjectives and verbs but not in nouns.

For verbs, the prediction was that the verbs would be learned in the (transitive) verb-CVC position since Verb-Object forms a more complete syntactic unit than does Subject-(transitive) Verb. This was the case in the Baker and McCarthy (1975) study. The results of the present study

indicated no such preference, and again the question to be answered is whether there was an equally facilitative effect or whether the semantic factors overrode the syntactic ones. One way of exploring this question is to consider it as a direct comparison between this study and the Baker and McCarthy study.

The experimental procedures used in the two studies were nearly identical, the only major difference being the randomized pairings of the real words with the CVC's in the present study (see Chapter Two). It remains, therefore, to look at the differences between the two sets of words.

One difference between the two sets of words was that the words chosen for the Baker and McCarthy study were all of high frequency. The words chosen for the present study represented a wide range of frequency of occurrence, but none were as common as all the words in the previous study. Why this would afect their learnability is not immediately clear especially when it is recalled that the analysis revealed no influence attributable to frequency per se.

It was suggested earlier in the discussion of the problems associated with using traditional word counts as indices of frequency that the high-frequency words used by Baker and McCarthy might be immune to these problems. The implication of that observation for the comparison being drawn is that for very common words, subjects had little

reason to pay attention to the semantic properties of the words because they were so familiar with them. Instead of attending to the semantic properties, they attended primarily to the syntactic properties of the words. In the present study, because the words varied in familiarity to the subjects, the subjects might have attended more to the semantic properties. They are obviously a more "interesting" set of words. It is a reasonable assumption that as lexical items become less familiar to a subject, he must concentrate more on their meaning attributes in order to remember them and, in this task, to bond them to the CVCs. Possibly, then, subjects in this study differentiated and associated the experimental words on the basis of their semantic characteristics and ignored their syntactic bonding properties. (Additional evidence for this explanation will be given in the discussion of the analysis of covariance later in this chapter.)

Even though the problem was to learn the responses, i.e., the nonsense syllables rather than the stimuli, the properties of the stimuli must be considered as a primary source of variation in the learnability of the paired associates. This is true for two reasons. In the first place, the nonsense syllables were judged to be of equal learnability (Glaze, 1927), and for each subject they were randomly paired with the stimuli to eliminate any possiblity of unique bonding. In the second place, the task required

of the subjects was not only to learn the response but to bond it with its correct stimulus. Since the responses were controlled for semantic content and the pairings randomized for each presentation, it is reasonable to assume that the semantic properties of the stimuli exerted some influence on the bonding phase of the task. While it is difficult to find support in the relevant literature for the argument that frequency may influence, or interfere with, syntactic facilitation, it should be noted that Pavio (1963) expressed a concern that high frequency words, because they exhibited only small differences in abstractness, might confound experimental results. Since the strength of the correlation between abstractness and imagery suggests that the two might measure the same underlying variable, it might be argued that <u>if</u> high frequency words interact with abstractness to confound experimental results, they might also interact with imagery with the same consequence.

Some support for the preceding explanation comes from post-experiment interviews with some of the subjects. Several reported that despite the fact that they saw half the stimulus items in each position, they tended to ignore the position of the stimulus items entirely. In effect, this would cause them to form associations on the basis of semantic properties and ignore syntactic category. More accurately, even though they sometimes formed syntactic units, they managed to do so whatever the position of the

stimulus. One subject reported, for example, that for the pair, dax expected, he recalled expected tax (return). He disregarded the position of the stimulus on the screen, possibly attending to temporal rather than spatial order. Another subject reported that he formed verb + object constructions for all the verbs. When asked why he did not form subject + verb constructions for those verbs presented on the left, he replied that he had not noticed that half of them were on the left. It appears, in short, that some subjects were able to ignore the position of the stimulus despite the experimenter's efforts to force them to attend to it and to read from left to right on each presentation.

In brief, the low frequency of the verbs (and possibly the nouns), relative to the ones used by Baker and McCarthy, may have contributed to the apparent increased salience of the semantic factors. Adjectives, as shown earlier, seemed slightly less influenced by the semantic factors (recall that the correlation between m and learnability for adjectives was .69 compared to .81 for nouns) but were influenced by position relative to the CVC, suggesting that the constituent integrity of Adjective-Noun has some psychological salience.

Table 6 shows the results of the covariance analysis adjusting the data for meaningfulness. There is a 39% reduction in the F-ratio for syntactic category and 88% reduction in the F-ratio for imagery. The latter is not a

surprising finding considering the results of the earlier analyses, but the reduction in the mean square for syntactic category is interesting. It suggests that meaningfulness plays a different role for some syntactic categories than for others.

It was suggested earlier that the adjectives were the aberrant class, specifically because of the unusually high m value for bad. Removing bad from considerationn raised the correlation between m and learnability from .44 to .69. Furthermore, it was the adjectives which revealed the only evidence for syntactic facilitation. A reasonable hypothesis at this point is that as syntactic category increases in importance, the importance of the semantic factors, in this case, meaningfulness, decreases. This is borne out by the correlations presented earlier and provides a reasonable explanation for the results of the covariance analysis.

Table 7 shows the summary of the ANOVA performed on the same factors as listed above but with response latency as the learning measure. As in the ANOVA with percent correct as the underlying learning measure, there were significant effects attributable to syntactic category and imagery. The syntactic category by position interaction was, however, non-significant.

TABLE 7

ANALYSIS OF VARIANCE SUMMARY TABLE

(Response Latency)

Source	Sum of Squares	Degrees of Freedom	Mean Squares	F Ratio	Probability
Category (C)	0.273	2.	0.137	10.604	0.001**
Imagery (I)	0.567	1.	0.567	43.999	0.001**
C x I	0.065	2.	0.032	2.504	0.099
Within Subjects	0.387	30.	0.013		
Position (P)	0.001	1.	0.001	0.107	0.745
C x P	0.016	2.	0.008	1.550	0.229
I x P	0.0	1.	0.0	0.0	0.999
C x I x P	0.015	2.	0.008	1.487	0.242
C x Within Subjects	0.153	30.	0.005		

CHAPTER IV

DISCUSSION

The present study attempted to investigate further the effects of syntactic facilitation on paired-associate learning while at the same time examining the effects of certain non-syntactic variables, i.e., frequency, abstractness, meaningfulness, and imagery. It was expected that the results of the Baker and McCarthy (1975) study with regard to syntactic facilitation would be confirmed. Specifically, it was expected that adjectives would show a "position" effect, the preferred learning order being Adjective-CVC, and that verbs would show a similar effect, the preferred order being (transitive) Verb-CVC. Based on the findings of Paivio and others (reviewed in Chapter I), the experimenter also expected to find that rated imagery was a highly influential variable in measuring the learnability of the words. It was also expected that abstractness, meaningfulness, and frequency would influence learning as well, though the degree of influence was predicted to be low relative to imagery. The results of this study must be interpreted and evaluated, then, on their contribution to either the confirmation or falsification of the hypotheses, on their contribution to knowledge about paired-associate learning, and, in a larger sense, on their

contribution to the sciences of linguistics and psychology.

The Baker and McCarthy data suggested that in the PA learning task, subjects took advantage of syntactic information when learning adjectives and verbs, but that for the nouns it was impossible to determine whether there was no syntactic facilitation or equal facilitation for the words in either position. The results of the present study replicated their findings with regard to adjectives and nouns, but found no evidence of syntactic facilitation in the learning of the verb lists.

Certainly it is in the adjectives where the strongest evidence for syntactic facilitation would be expected, but one would also expect to find some evidence in the verb lists. All the verbs used were unambiguously transitive. A reasonable prediction, based on traditional descriptive phrase-structure, was that the Verb-CVC pattern would be easier to learn than the CVC-Verb pattern. The transitive verb-object relationship is, after all, a closer one, in most grammatical descriptions, than the subject-verb relationship. In transformational grammars, for example, both the transitive verb and its object are immediately dominated by the Verb Phrase node while the subject noun phrase is immediately dominated by the Sentence node. It must be emphasized once more, however, that there is no *a priori* reason for expecting subjects' behavior in this task to correspond to any particular formal grammar *except* that

the results of Baker and McCarthy's study were interpreted in such a way as to suggest correspondence between their subjects' behavior and traditional phrase structure.

A second reason for expecting to find a Verb-CVC order preference comes from examination of the data collected by Baker and Hull (1975) on meaningfulness. The m ratings given in Table 1 and referred to throughout the analysis, represent the mean number of responses to any given word, but the raw data for that study reveal that the nature of the responses provides much valuable information as well. It was found, for example, that the preferred category for associates to verb stimuli was nouns: 58% of all responses given to verb stimuli were nouns, compared with verb responses which accounted for 27% of the total. This contradicts Ervin-Tripp's finding (1970, reported in Chapter I) that adults tend to give paradigmatic responses rather than syntagmatic responses in word association tasks.

In fairness, it must be admitted that Ervin-Tripp's conclusions were based on nouns only, and that Baker and Hull's results for nouns were in accord with those conclusions (68% of the responses to nouns were nouns, 21% were adjectives, and only 6% were verbs). But the relevant information concerns verbs. Not only were the associates overwhelmingly nouns, but inspection of the associates revealed that they were almost without exception nouns that could logically serve as objects, e.g., worm, frog, and cat,

given for dissected, rather than subjects, although there were some locative or "general" responses, e.g., laboratory or biology. This is not surprising, of course, since there is a highly limited set of nouns in the language, specifically those which are, in feature terms, [+animate] or, for some verbs, [+human], which can function as agents for transtive verbs. The nouns which can serve as objects, on the other hand, come from a far wider class of nouns. Ideas, arguments, and bananas, for example, can all be dissected as well as worms and frogs, but none of these can do the dissecting. The preponderance of object-type responses to the verb stimuli suggested that in the PA learning task, subjects might learn the Verb-CVC more easily than the CVC-Verb order. This was precisely the case in Baker and McCarthy's (1975) study, but not in the present study. The question that arises from this discrepancy is, simply, why did it occur?

It was suggested in the previous chapter that a contributing factor might have been the difference in the degree of familiarity to subjects of the words in the two studies. Specifically, the hypothesis was that because the words in the present study were less familiar than Baker and McCarthy's words, subjects found them more interesting semantically and focused on their semantic rather than their syntactic properties. The subjects, in other words, might have been able to ignore the position of the real word

relative to the CVC and concentrate on its semantic properties. In making "sense" of the nonsense syllable paired with dissect, then, subjects might have used the same strategies for either order since dissect has relatively high semantic interest. This could explain the results for nouns and verbs although there is at least one alternative possibility for nouns. Baker and Hull's (1976) data showed that the favored response category for noun stimuli in an association task is nouns. The fact that only nouns elicit paradigmatic responses suggests that in the PA task, subjects might have been seeing the CVC's paired with nouns also as nouns. If that were the case, there should be no position preference.

The question remaining to be answered, though, is this: If the relative unfamiliarity of the items in the present study influenced subjects' learning in a particular way for verbs (and possibly nouns), why did it not influence their learning of adjectives? Of course, it is possible that it did, to a lesser degree, but the fact is that the adjectives demonstrated a position effect, i.e., were easier learned when positioned to the right of the CVC, when the other two syntactic categories did not. A reasonable explanation is that the unit integrity of the adjective-noun in English is sufficiently strong to overcome obstacles which block the emergence of a word-order effect in weaker syntactic units. In contrast to Paivio's (1963) assertion that well-

established syntactic patterns can be overridden, the point being made here is that a sufficiently well-established syntactic configuration might override semantic attributes in importance or salience to the subject.

Strong support for this position comes from the Baker and Hull study on meaningfulness cited earlier. They found that the associative responses given for adjective stimuli were overwhelmingly nouns. Of the total responses given to 12 adjective stimuli, 56% were nouns while only 35% were adjectives, and 9% were from other syntactic categories. These findings, in combination with the finding of a position effect discussed earlier, make it possible to make a fairly strong claim about the psychological salience of this particular construction. If the bonding between the adjective and nonsense form were made purely on the basis of meaning, i.e., if the unit integrity of Adjective-Noun were based purely on the fact that adjectives and nouns are bound by virtue of their semantic association (and there is nothing in the meaningfulness data to suggest otherwise), then it would not be expected that subjects would show any position preference. Still, a cautionary note must be sounded. It would be a mistake to generalize this conclusion to other syntactic units without first establishing that they behave in the same way under experimental conditions. It seems at this point that the strongest claim that can be made is that linguists have described, in this particular

case, a formal constituent that can be shown to have psychological import.

In addition to the hypothesis about syntactic facilitation, a second hypothesis was under consideration, namely that rated imagery would dominate the other semantic factors in measuring learnability. The results with regard to this hypothesis were perhaps the most interesting of the experiment. They will be discussed in terms of Paivio's "conceptual peg" hypothesis.

Paivio (1963, 1969) suggested that a syntactic pattern might be overridden when semantic attributes of the noun were very strong. In particular, he suggested that a stimulus member of a pair serves as a "conceptual peg" on which its associate is "hung." Nouns would, he reasoned, serve as superior pegs because of their higher rated imagery. There are two implications of the present data for Paivio's hypothesis. The hypothesis itself receives weak support from the fact that there was a significant effect due to syntactic category in the ANOVA and that both learning measures showed nouns easier to learn. On the other hand, there is no evidence to suggest that it is the relatively high rated imagery of nouns which is contributing to this effect. In the first place, both adjective and verb lists contained high as well as low imagery items (see Table 1). Second, there was no significant interaction between imagery and syntactic category either before or after the

adjustment for meaningfulness. Even more significantly, the covariance analysis adjusting the data for meaningfulness, showed a 39% reduction in the F-ratio for syntactic category and an 88% reduction for imagery. This result provides strong support for the claim that the factor which most influenced learning was meaningfulness and not imagery as has been suggested (Paivio, 1963).

There are two reasons why imagery has been traditionally emphasized over meaningfulness as heavily influencing learning in experiments similar to the one reported here. The first is that conclusions about the effect of imagery have been based largely on studies which have examined only the effect of scaled imagery, i.e., quite independent of any other semantic properties. Similarly, there is a preponderance of literature measuring the effect of meaningfulness (mostly of nonsense forms) on verbal learning, but there have been relatively few attempts made to investigate simultaneously the effects of meaningfulness and imagery on the learning of real words. It has been, therefore, impossible to make any sensible judgments about their independent effects on learning.

The second reason has to do with the relative naivete of the analytic techniques used. Most researchers have been content to interpret a high correlation between I and the learning measure or a significant F-ratio in an analysis of variance as sufficient evidence for claiming the superior

status of imagery as a predictor of learnability. Indeed, if the analysis of the present data had stopped with the Pearson and Spearman correlations, the evidence for the claims being made would have been considerably weaker. The partial correlations computed on the data showed that the influence of imagery on the learning measures was largely dependent upon their mutual correlation with meaningfulness (as described in Chapter III). The analysis of covariance also provided a more sensitive indicator of the effect of a variable since it permits any other variable to be held constant. The analysis in the previous chapter showed, for example, an 88% reduction in the variance for imagery when meaningfulness was held constant.

It seems, in short, that psychologists may have been looking at the wrong variable, and that their conclusions should be reconsidered in light of the results reported here. Furthermore, future experimentation should look at the two properties simultaneously and the analysis of the data should utilize more sensitive techniques such as partial correlation and analysis of covariance.

The results of the present study provide some interesting commentary on other experiments reported in the first chapter. Broadly, the results seem to suggest once more that descriptive phrase structure does not necessarily correspond to the phrase structure assigned an utterance by the user. Martin (1970) found this to be the case with

sentences. While the present study was limited to paired associates and, thus, may not generalize to sentential processing, it seems safe to acknowledge at least that the chunking of linguistic material is not predictable from the syntactic structure that can be externally imposed.

The point which emerged from the Martin and Levelt studies, i.e., that there is no a priori reason for assuming that any particular description (grammar) of the language product should correspond to any kind of language processing, was made again, even though the results suggested that the minor constituent Adjective-Noun might have some psychological salience.

The results of the present study suggest that Epstein (1961) might have been slightly off center in his explanation for subjects' ability to learn syntactically structured nonsense material more easily than unstructured material. He concluded that the learner took advantage of syntactic structure when it was available to him. It might be more precise to say that the learner used syntactic information because it was the only indisputable information available to him (dealing with nonsense "sentences"). The real words in the present study had several semantic attributes attached to them. It may have been the case that subjects simply did not need the syntactic information, that they were unaware of the syntactic information because they were able to construct strong semantic bonds independent of

English syntax except in the case of the adjectives where the English syntactic pattern may be firmly enough established that it influenced subjects' learning of the pairs. This is a similar point to the one made in the comparison of Baker and McCarthy's findings with the present ones with regard to verbs. It was suggested that as the semantic information becomes less "interesting," the syntactic information increases in relevance. In Epstein's study, there was little, if any, semantic information that could be extracted from the nonsense items until they were organized into "sentences, thereby providing each item with a syntactic "label," and making them easier to learn. Similarly, in the Baker and McCarthy study, because all the words were extremely common, and less interesting semantically, subjects might have been forced to attend to their syntactic properties.

The findings of the present study are interesting with regard to the Glanzer (1962) experiment. First of all, there seems to be little doubt that Baker and McCarthy (1975) were correct in their criticism of Glanzer's assumption of content class homogeneity. There are clearly learnability differences among the three syntactic categories, nouns, verbs, and adjectives, which Glanzer combined in his content classes. These differences, present in Glanzer's data but pointed out by Baker and McCarthy, are also present in the results reported here.

In the second place, these findings differ from Glanzer's in that he found no "position by type" interaction. In his first experiment, he found that all word types were learned better when they followed the nonsense items. There was simply no difference between syntactic categories which would have been expected had syntactic facilitation been operative. The present analysis did reveal a significant type by position interaction, showing that subjects apparently did use some kind of syntactic facilitation when learning the adjective-CVC pairs.

In comparing Glanzer's (1962) and Kanungo's (1969) findings that CVC-Word order was more easily learned than Word-CVC order with the present data, it can be seen that the latter may have resulted from a more sensitive experimental design. Both Glanzer and Kanungo sabotaged their results by using CVC's as stimuli for half their subjects and real words as stimuli for the other half. The objections to this technique were detailed in Chapter I. The present experiment did reveal a position by type interaction attributable to the adjectives. As English syntax would predict, learning was better for the Adjective- CVC pattern than for the CVC-Adjective pattern.

The results reported here must be considered in terms of their contribution to paired-associate learning, and more importantly, on how these findings contribute to

understanding of that hazy and rather poorly defined area where the sciences of linguistics and psychology meet.

The effect of meaningfulness on PA learning was an especially interesting aspect of this study. Since meaningfulness values had not been obtained for adjectives and verbs before Baker and Hull (1976), there were no studies available to document the effect of meaningfulness on the learning of these syntactic categories comparable to the Paivio studies with nouns. Furthermore, previous studies failed to control other semantic factors as closely as the present one, so there was much to be learned about the effect of (stimulus) meaningfulness on paired associate learning.

As has been indicated repeatedly, meaningfulness exerted more influence on the learning of the paired associates than any of the other properties under consideration. The absence of a significant interaction effect between syntactic category and imagery both in the original analysis of variance and after the data were adjusted for meaningfulness suggests that the effect was roughly equal across syntactic categories. This offers support for claims about the influence of meaningfulness in PA learning and bolsters confidence in the utility of m as a measure of one of the many dimensions of meaning. The richness of association of a word, measured by the procedure described earlier, clearly influences the learnability of

the word in a PA task. It was demonstrated that the influence of imagery was dependent on its relationship to meaningfulness. This is a particularly interesting point since it has traditionally been believed that meaningfulness had greater effect on the response member of the pair while imagery had greater influence on the stimulus (see Chapter I). In the present study, of course, the responses were always nonsense syllables, so the effects of imagery and meaningfulness can only be discussed in terms of stimuli. The point is, however, that the traditionally held views must be met with some skepticism if it holds in general that imagery exerts its influence via meaningfulness. One cautionary note must be sounded: It would be a mistake to generalize the conclusions drawn here to all words which might be chosen for use in a PA learning task. The suggestion was made earlier that the relationship between imagery and meaningfulness and between abstractness and meaningfulness might differ for words with highly emotional or evaluative connotations.

Although the presence of a significant position by syntactic category interaction effect in the ANOVA makes any interpretation of the main effect of syntactic category somewhat tenuous, it is interesting to note that averaging over both positions, the order of learnability for the three syntactic categories was noun, adjective, verb. The nouns and adjectives were very close at 70 and 68.8 (arc sine

transformation of percent correct) with verbs lagging behind at 65. An analysis of simple effects showed the verbs to be significantly different from both the nouns and adjectives in learnability when presented on either the left or right of the CVC. There was no significant difference between nouns and adjectives on the right, but on the left, there was a difference, the adjectives being significantly easier to learn.

The question which now arises is what can the information gained about PA learning contribute to the understanding of larger issues in psycholinguistics. In other words, of what concern to linguists and psychologists are the results that have been reported here?

To answer this question, it is necessary to define the domain of inquiry, or at least to narrow it down. If the domain were formal, descriptive linguistics, then the answer would be simple, straightforward, and negative. But linguistics in the second half of the twentieth century has moved away from well-established descriptivism into an area that was once strictly the domain of psychologists. The intricacies of the interrelationships between language and mind are being explored in their many aspects: memory, language development, and cognition, to name but a few.

As suggested in Chapter I, one area of research which has been pursued by both linguists and psychologists is the

nature of linguistic units larger than words. Definitions or specifications of these units were vague on both sides, and this vagueness and confusion about each other's views led in part to a great deal of research by linguists seeking to establish "psychological reality" for their formal, descriptive rules. That these endeavours were largely unsuccessful is well documented in the linguistic literature of the past two decades, and several specific instances were reported in Chapter I. The reasons for their failure are patently obvious and need no elaboration here except to state that the central assumptions, i.e., that descriptive phrase structure, or indeed, rules of any kind, might correspond to psychological processes, was naive.

The present research undertook an investigation of the nature of these "units" or "chunks," the existence of which has generally been accepted by both linguists and psychologists. In a sense, though, its goals were more specific. Since the experiment was carefully constructed to integrate semantic, or meaning, properties and syntactic, or formal, properties, what was really being explored was the relationship between semantics and syntax in the chunking, or bonding, process. It is in this area that the research should contribute to psycholinguistic science.

Perhaps one objection should be anticipated and answered before addressing the issue. It concerns generalizing from a paired-associate paradigm to sentential

processing from data gathered on the learning of paired associates. That people take advantage of semantic and, possibly to a more limited degree, syntactic information in the decoding of sentences cannot be argued. But to restrict an investigation of the relationship between meaning and form to the study of sentences would be to introduce so much experimental "noise" as to make precision of analysis and interpretation extremely difficult. From the point of view of experimental control, there is much to said for the PA paridigm, which permits a high degree of control, as a method for the simultaneous investigation of the semantic and syntactic properties that influence bonding.

Once the validity of the paradigm is accepted, the substantive question of the relationship between semantics and syntax can be addressed. Baker, Prideaux, and Derwing (1973) pointed out that pure syntax, i.e., totally devoid of any semantic considerations, does not exist outside formal analytical systems. That point is well documented in the literature and is relevant for the present discussion in that it emphasizes the futility of attempting to analyze, interpret, or discuss syntactic properties independent of semantics. The authors went on to distinguish two different aspects of semantics, one which refers to the sentence content in terms of the specific lexical items and the functional relationships between them, and one which refers to the "implications of syntactic markers or, more

generally, the semantic significance of a specific syntactic pattern" (p.28).

This is a particularly useful distinction for understanding the relevance of the present research. It might be reasonable to argue that a substantial amount of information has been gained with regard to both aspects of semantics.

Some light is shed on the aspect of semantics dealing with the significance of particular syntactic patterns by the fact that subjects apparently attended to one syntactic pattern and ignored others. More precisely, the Adjective-Noun syntactic pattern seemed to have more semantic significance than any other pattern that was deemed possible in the paradigm. Perhaps even more important are the implications of the nouns and verbs not showing any evidence of syntactic facilitation. That fact may suggest that other factors override established syntactic patterns, or it may suggest - and this point is especially important for linguists - that the syntactic patterns are as yet insufficiently defined. Baker and McCarthy's (1975) suggestion that equal facilitation may have occurred for nouns implies that the delineation of syntactic patterns in terms of the speaker/hearer's "awareness" of them, might be inadequate. If the patterns cannot be or are inaccurately defined, they cannot be reasonably tested, or said either to exist or not to exist. In other words, it is premature to

assume that syntactic patterns do not influence the learning of nouns and verbs, or that nouns and verbs do not bond somehow with other elements to form syntactic constituents. Rather, it is more reasonable to assume tha certain syntactic patterns that are currently thought to have some psychological validity, e.g., verb-object, are of little importance in the learning task and that a re-specification of syntactic pattern might yield more fruitful results.

The research reported here is of even greater interest with regard to the second aspect of semantics, i.e., content in terms of particular lexical items and the functional relationships between them. In normal connected discourse, the speaker/hearer takes advantage of both aspects of semantics described above. Even in the experiment described here, he had limited amounts of information about both aspects. He knew the syntactic category of the word and the "meaning" of the word itself. He was not explicitly given the functional relationship between the word and the CVC, but neither was he prohibited from creating or "seeing" relationships. He was, in fact, by the design of the experiment, encouraged to do so. It was the nature of these relationships that proved to be interesting.

It is clear that the overriding influence was semantic in the sense of Baker, Prideaux, and Derwing's first aspect of semantics, i.e., the one referring to the specific lexical items and the functional relationships between them.

Specifically, it was discovered that these functional relationships are not necessarily defined by the relationships that have been labelled by linguists, no matter the particular formal system, as syntactic. More important was the richness of association of the particular items, a dimension of meaning closely related to past experience, non-linguistic as well as linguistic. It was found, furthermore, that the imageability of a lexical item contributed to the bond it formed with another item, but that imageability was very clearly related to the richness of association.

All this leads to the inescapable conclusion that since language is essentially a meaningful activity, subjects in experimental situations respond to language stimuli in terms of possible meaning, and not in terms of syntactic form <u>even when the meaning is minimized</u> by presenting the stimuli in isolation, i.e., in context-free environments. It is suggested that linguists in their formulation of grammatical descriptions should be sensitive to the fact that the relationships between lexical items are primarily a function of their semantic content. It is suggested, moreover, that psychologists should remember - since in fairness it cannot be claimed that they have not begun to realize - that formal grammars do not necessarily provide fruitful, or even particularly interesting, paradigms for research into language processing.

Summary

This study investigated the psychological salience of English constituent structure in the learning of real word-CVC paired associates. At the same time it investigated the effects of the semantic parameters, frequency, meaningfulness, imagery, and abstractness. The research was intended, in a larger sense, to shed light upon the nature of an aspect of language processing commonly called "chunking."

It was discovered that syntactic facilitation was operative in the learning of adjective-CVC pairs, but that it was impossible to determine whether or not it was operative for nouns or verbs paired with CVC's. Furthermore, it was found that meaningfulness exerted more influence on the learning task than any of the other factors. This finding was interesting in light of previous studies which have found imagery (and the highly correlated factor, abstractness) to be highly influential. In addition, the effect of frequency measured by traditional word counts was found to be insignificant, and a more sensitive measure suggested.

The data suggest that the relationship between have traditionally been classified as semantic factors and factors commonly called syntactic is an intricate and

variable one. Certainly it is impossible to claim the superiority of one over the other in the chunking process. It can be said with a fairly high degree of certainty, however, that bonding entails the use of information both about syntactic patterns in the language and about particular meaning dimensions of the lexical items. Furthermore, it seems to be the case that a strong semantic factor can override a strong syntactic factor, and, conversely, that a sufficiently well established syntactic pattern can maintain its influence even in the presence of strong semantic factors.

BIBLIOGRAPHY

Anderson, John and Bower, Gordon, H. Human associative memory. Washington: V.H. Winston and Sons, 1973.

Archer, E.J. A re-evaluation of the meaningfulness of all possible CVC trigrams. Psychological Monographs, 1960, 74.

Bacharach, V.R., Kellas, George, and McFarland, Carl. Structural properties of transitive and intransitive verbs. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 486-490.

Baker, W.J., Prideaux, G.D., and Derwing, B.L. Grammatical properties of sentences as a basis for concept formation. Journal of Psycholinguistic Research, 1973, 2 (3).

Baker, W.J. and Hull, T. Imagery, abstractness, and meaningfulness norms for 100 nouns, adjectives, and verbs. Unpublished manuscript. The University of Alberta, 1976.

Baker, W.J. and McCarthy, Karel L. Word category effects on paired-associate learning. Unpublished manuscript submitted in lieu of thesis, University of Alberta, 1975.

Blough, D.S. and Lipsitt, L.P. The discriminative control of behavior. In J.W. Kling and L.A. Riggs (Eds.), Woodworth and Schlosberg's experimental psychology, (3rd. ed.). New York: Holt, Rinehart and Winston, 1971.

Bowers, Henry. Memory and mental imagery. British Journal of Psychology, 1931, 271-282.

Brown, R., and Berko, J. Word association and the acquisition of grammar. Child Development, 1960, 31, 1-14.

Butler, B.E. and Merikle, P.M. Uncertainty and meaningfulness in paired-associate learning. Journal of Verbal Learning and Verbal Behavior, 1970 9, 634-641.

Butter, M.J. and Palermo, D.S. Effects of imagery on paired-associate recall as a function of retention interval, list length, and trials. Journal of Verbal Learning and Verbal Behavior, 1970,)u 9, 716-719.

Carroll, J.B., Davies, P., and Richman, B. The American heritage word frequency book. Boston: Houghton Mifflin, 1971.

Cairns, H. S. and Cairns, Charles E. Psycholinguistics, a cognitive view of language. New York: Holt, Rinehart,and Winston, 1976.

Chomsky, Noam, Syntactic structures. The Hague: Mouton, 1957.

Chomsky, Noam. Formal proprties of grammars. In R.D. Luce, R. Bush, and E. Galanter (Eds.), Handbook of mathematical psychology, Vol. 11, 323-418, New York: Wiley, 1963.

Chomsky, Noam. Aspects of the theory of syntax. Cambridge, Mass.: The M.I.T. Press, 1965.

Cofer, Charles N. Properties of verbal materials and verbal learning. Chapter 19 in J.W. Kling and Lorin Riggs (Eds.), Woodworth and Schlosberg's experimental psychology. New York: Holt, Rinehart and Winston, 1971.

Darley, F.L, Sherman, D. and Siegel, G.M. Scaling of abstraction level of single words. Journal of Speech and Hearing Research, 1959, 2 (2), 161-167.

Delprato, D. J. and Baker, E.J. Concreteness of peg words in the mnemonic systems. Journal of Experimental Psychology, 1974, 102 (3), 520-522.

Dominowski, R.L. and Gadlin, H. Imagery and paired-associate learning. Canadian Journal of Psychology, 1968, 22, 336-348.

Ellis, Henry. Fundamentals of human learning and cognition. Dubuque, Iowa: W.M.C. Brown Co., 1972.

Epstein, W. The influence of syntactic structure on learning. American Journal of Psychology, 1961, 74, 80-85.

Epstein, W. Backward association as a function of meaningfulness. Journal of General Psychology, 1962, 67, 11-20.

Ervin-Tripp, Susan. Changes with age in the verbal determinants of word-association. In L. Jakobovits. and M. Miron (Eds.), Readings in the psychology of language. Englewood-Cliffs, New Jersey: Prentice Hall, 1967.

Ferguson, G.A. Statistical analysis in psychology and education, third edition. New York: Mc-Graw Hill, 1971.

Fodor, J.A. and Bever, T.G. The psychological reality of linguistic segments. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 414-420.

Fodor, J.A., Bever, T.G., and Garrett, M.F. The Psychology of Language. New York: McGraw-Hill, 1974.

Garrett, H.E. Statistics in psychology and education. New York: Longmans, Green, and Co., 1958.

Garrett, M., Bever, T.G., and Fodor, J.A. The active use of grammar in speech perception. Perception and Psychophysics, 1966, 1, 30-32.

Glanzer, M. Grammatical category: a rote learning and word association analysis. Journal of Verbal Learning and Verbal Behavior, 1960, 1, 31-41.

Glaze, J.A. The association value of nonsense syllables. Journal of Genetic Psychology, 1928, 35, 255-267.

Gorman, A.M. Recognition memory for nouns as a function of abstractness and frequency. Journal of Experimental Psychology, 1961, 61, 23-29.

Goss, A.E. Comments on Professor Noble's paper. In C.N. Cofer and B.S. Musgrave (Eds.), Verbal behavior and learning: Problems and processes. New York: McGraw-Hill, 1963.

Goss, A.E., and Nodine, C.F. Paired associates learning: the role of meaningfulness, similarity, and familiarization. New York: Academic Press, 1965.

Haagen, C.H. Synonymity, vividness, familiarity, and association value ratings of 400 pairs of common adjectives. Journal of Experimental Psychology, 1949, 27, 453-463.

Hayes, J.R.M. Memory span for several vocabularies as a function of vocabulary size. Quarterly Progress Report. Cambridge, Mass.: Acoustics Laboratory, M.I.T., January-June, 1952.

Hilgard, E.R. Methods and procedures in the study of learning. Chapter 15 in S.S. Stevens (Ed.), Handbook of experimental psychology. New York: Wiley, 1951.

Herriot, Peter. Attributes of memory. London: Methuen and Co., Ltd., 1974.

Hulicka, I.M. and Grossman, J.L. Age-group comparisons for the use of mediators in paired-associate learning. Journal of Gerontology, 1967, 21, 46-51.

Hull, C.L. The meaningfulness of 320 selected nonsense syllables. American Journal of Psychology, 1933, 45 730-734.

Jenkins, J.J., Russel, W.A. and Suci, G. An atlas of semantic profiles for 360 words. American Journal of Psychology, 1958, 71, 688-699.

Johnson, N.F. Linguistic models and functional units of language behavior. In S. Rosenberg (ed.)< Directions in psycholinguistics. New York: Macmillan, 1965. Pp. 29-65.

Johnson, S.C. Hierarchical clustering schemes. Psychometrika, 1967, 32 241-254.

Kanungo, R. Learning of content, function, and nonsense words: an extension of Glanzer's experiment. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 262-265.

Keenan, Edward. Universals of passives reconsidered. Lecture given at Calgary, Alberta, March 18, 1978.

Kintsch, W. Abstract nouns: imagery versus lexical complexity. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 59-65.

Kling, J.W. and Riggs, L.A. (Eds.) Woodworth and Schlosberg's experimental psychology, (3rd. ed.). New York: Holt, Rinehart and Winston, 1971.

Krueger, W.C.F. The relative difficulty of nonsense syllables. Journal of Experimental Psychology, 1934, 17, 145-153.

Lambert, Wallace. Associational fluency as a function of stimulus abstractness. Canadian Journal of Psychology, 1955, 9 (2), 103 106.

Levelt, W.J.M. The perception of syntactic structure. Heymans bulletin No. HB-69-30EX. Groningen University Department of Psychology, Groningen, Netherlands, 1969a.

Levelt, W.J.M. The scaling of syntactic relatedness: a new method in psycholinguistic research. Psychonomic Science, 1969b, 17, (6), 351-352.

Levelt, W.J.M. A scaling approach to the study of syntactic relations. In G.B.F. D'Arcais and W.J.M. Levelt (Eds.) Advances in psycholinguistics. Amsterdam: North-Holland, 1970.

Levelt, W.J.M. Hierarchial chunking in sentence processing. Perception and Psychophysics, 1970, 8, (2).

Mandler, G. and Huttenlocher, J. The relationship between associative frequency, associative ability and paired-associate learning. American Journal of Psychology, 1956, 69, 424-428.

Martin, Edwin. Toward an analysis of subjective phrase structure. Psychological Bulletin, 1970, 3, 153-166.

Martin, E. and Roberts, L.H. Grammatical factors in sentence retention. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 211-218.

McGehee, Nan and Schulz, R. Mediation in paired-associate learning. In Readings in the psychology of language. Englewood Cliffs, New Jersey: Prentice Hall, 1967.

McGuire, W.J. A multi-process model for paired-associate learning. Journal of Experimental Psychology, 1961, 62, 335-347.

Miller, G.A. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 1956, 63, (2).

Miller, G.A. Some psychological studies of grammar. American Psychologist, 1962, 17, 748-762.

Miller, G.A. Psycholinguistic approaches to the study of communication. In D.L. Arm (Ed.), Journeys in science. Albequerque: The University of New Mexico Press, 1967.

Miller, G.A. A psychological method to investigate verbal concepts. Journal of Mathematical Psychology, 1969, 6, 169-191.

Millward, R.B. Theoretical and experimental approaches to human learning. Chapter 20 in J.W. Kling and L.A. Riggs (Eds.), Woodworth and Schlosberg's experimental psychology, (3rd. ed.). New York: Holt, Rinehart and Winston, 1972.

Nelson, D.L,, Brooks, D.H., and Wheeler, J.W. Sensory and meaning features in stimulus recognition and associative retrieval. Journal of Experimental Psychology, 1975, 104 (1, No.6), 711-719.

Noble, C.E. The role of stimulus meaning (m) in serial verbal learning. Journal of Experimental Psychology, 1952, 43, 437-446.

Noble, C.E. Meaningfulness and familiarity. In C.N. Cofer and B.S. Musgrave (Eds.), Verbal behavior and learning: Problems and processes. New York: McGraw-Hill, 1963. Pp. 119-155.

Noble, C.E. and McNeely, D.A. The role of meaningfulness in paired-associate verbal learning. Journal of Experimental Linguistics, 1957, 53 (1), 16-22.

Noble, C.E., Stockwell, F.E., and Pryer, M.W. Meaningfulness (m') and association value in paired associate syllable learning. Psychological Reports, 1957, 3, 441-452.

Noble, C.E. and Parker, G.V.C. The Montana scale of meaningfulness. Psychological Reports, 1960, 7, 325-331.

Osgood, C.E. Method and theory in experimental psychology. Oxford: Oxford University Press, 1953.

Osgood, C.E. On understanding and creating sentences. American Psychologist, 1963, 18, 735-751.

Paivio, Allan. Learning of adjective-noun paired associates as a function of adjective-noun word order and noun abstractness. Canadian Journal of Psychology, 1963, 370-379.

Paivio, Allan. Latency of verbal associations and imagery to noun stimuli as a function of abstractness and generality. Canadian Journal of Psychology, 1966, 20,, 378-387.

Paivio, A. A factor analytic study of word attributes and verbal learning. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 41-49.

Paivio, Allan. Mental imagery in associative learning and memory. Psychological Review, 1969, 76, (3), 241-263.

Paivio, A. and Foth, D. Imaginal and verbal mediators and noun concreteness in paired-associate learning: the elusive interaction. Journal of Verbal Learning and Verbal Behavior, (1970) 9, 384-390.

Paivio, A. and Olver, M. Denotative generality, imagery, and meaningfulness in paired-associate learning of nouns. Psychonomic Science, 1964, 1, 183-184.

Paivio, A. and Steeves, R. Relations between personal values and imagery and meaningfulness of value words. Perceptual and Motor Skills, 1967, 24, 357-358.

Paivio, Allan and Yuille, J.C. Mediation instructions and word attributes in paired-associate learning. Psychonomic Science, 1967, 8,, 65-66

Paivio, Allan, Yuille, J.C., and Madigan, S. Concreteness, imagery, and meaningfulness values for 925 nouns. Journal of Experimental Psychology Monograph, 1968, 76, 1-25.

Paivio, A. Yuille, J.C., and Smythe, P.C. Stimulus and response abstractness, imagery, and meaningfulness, and reported mediators in paired-associate learning. Canadian Journal of Psychology, 1966, 20, 362-377.

Paivio, Allan. Unpublished norms. University of Western Ontario, 1977.

Pate, J.L., Ward, P.and Harlan, K.B. Effects of word order and imagery on learning verbs and adverbs as paired associates. Journal of Experimental Psychology, 1974, 103, (4), 792-795.

Polzella, D.J. and Rohrman, N.L. Psychological aspects of transitive verbs. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 537-540.

Prytulak, L.S. Natural language mediation. Cognitive Psychology, 1971, 2, (1),.

Richardson, J.T.E. Imagery, concreteness, and lexical complexity. Quarterly Journal of Experimental Psychology, 1975, 27, 211-223.

Richardson, J.T.E. Concreteness and imageability. Quarterly journal of Experimental Psychology, 1975, 27, 235-249.

Rohrman, Nicholas L. More on the recall of nominalizations. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 534-536.

Rommetveit, R. Words, meanings and messages. New York: Academic Press, 1968.

Schonpflug, W. and Vetler, G. Relations among characteristics of CVC trigrams. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 157-158.

Snedecor, G.W. and Cochran. W.G. Statistical methods. Ames, Iowa: The Iowa State University Press, 1967.

Spreen, Otfried and Schulz, Rudolph. Parameters of abstraction, meaningfulness, and prononciability for 329 nouns. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 459-468.

Stratton, R. Jacobus, K., and Brinley, B. Age of acquisition, imagery, familiarity and meaningfulness norms for 543 words. Behavior research methods and instrumentation, 1975, 7 (1), 1-6.

Taylor, Janet A. Meaning, frequency, and visual duration threshold. Journal of Experimental Psychology, 1958, 55, (4), 329-334.

Thorndike, E.L. and Lorge, I. The teacher's wordbook of 30,000 words. New York: Teacher's College, 1944.

Underwood, B.J. and Schulz, R.W. Meaningfulness and verbal learning. Chicago: Lippincott, 1960.

Wind, M. and Davidson, M. Facilitation of paired-associate learning by language context. Psychonomic Science, 1969, 15, 184-185.

Winer, B.J. Statistical principles in experimental design. New York: McGraw-Hill, 1971.

Winnick, W.A. and Kressel, K. Tachistoscopic recognition thresholds, paired-associate learning, and free recall as a function of abstractness-concreteness and word frequency. Journal of Experimental Psychology, 1965, 70, 163-168.

Witmer, L.R. The association value of three-place consonant syllables. Journal of Genetic Psychology, 1935, 47, 337-360.

Wulff, J.J. and Stolurow, L.M. The role of class-descriptive clues in paired-associates learning. Journal of Experimental Psychology, 1957, 53 (3). 199-206.

Yates, F.A. The art of memory. London: Routledge and Kegan Paul, 1966.

Yngve, V.H. A model and an hypothesis for language structure. Proceedings of the American Philosophical Society, 1960, 104, 444-466.

Yuille, J.C. Concreteness without imagery in p.a. learning. Psychonomic Science, 1968, II, 55-56.

APPENDIX

NONSENSE SYLLABLES

CVC	GLAZE ASSOCIATION VALUE
PAIRED WITH NOUNS:	
BEP	13
WUB	0
JAT	20
VUF	7
YIG	7
GID	7
MEZ	20
DEJ	20
KEX	13
HUC	7
ZIL	0
COH	27
PAIRED WITH ADJECTIVES:	
HEG	20
WUP	13
SIJ	0
ZOR	20
VUD	27
YAZ	7
KEB	7
NUX	7
BOF	7
GIK	13
RIH	27
TIV	20
PAIRED WITH VERBS:	
NIZ	27
MIB	7
WUG	13
LAJ	0
FEP	20
TOV	0
VOC	27
YOM	7
JIK	13
DAX	0
BUH	27
KEF	20